WO9707245

Publication Title:

METHODS AND DEVICES FOR PARALLEL MULTIPLEX POLYNUCLEOTIDE SEQUENCING

Abstract:

The present invention is directed to methods of multiplexing sequencing reactions such as the Sanger method when used with appropriately prepared complex samples. The method is advantageous because it allows the sequencing of many polynucleotides simultaneously. The method is primarily based on the Church and Gilbert multiplex sequencing reactions, in that the polynucleotides to be sequenced are constructed to contain terminal base-specific labeling. Furthermore, the method obviates the need to cross-link reaction products to membranes with chemical reagents. Method steps comprise, parallel separation of sequencing templates or sequencing reaction products from mixture prior to gel electrophoresis, and parallel detection of different samples or reaction products. To achieve separation, plural samples are applied to specific separation channels by parallel sample-specific binding means to said specific channels, thereby combining both multiplexation and parallelism. Detection is achieved by detector means that comprises tag-probe arrays. Suitable devices for performing the method are also described.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, G01N 27/26, C07H 21/00	A1	(11) International Publication Number: WO 97/07245 (43) International Publication Date: 27 February 1997 (27.02.97)

(21) International Application Number: PCT/US96/13313

(22) International Filing Date: 14 August 1996 (14.08.96)

(30) Priority Data:
08/514,820 14 August 1995 (14.08.95) US

(71)(72) Applicant and Inventor: RABANI, Ely, Michael [US/US]; 4495 Vision Drive #1, San Diego, CA 92121-1942 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHODS AND DEVICES FOR PARALLEL MULTIPLEX POLYNUCLEOTIDE SEQUENCING

(57) Abstract

The present invention is directed to methods of multiplexing sequencing reactions such as the Sanger method when used with appropriately prepared complex samples. The method is advantageous because it allows the sequencing of many polynucleotides simultaneously. The method is primarily based on the Church and Gilbert multiplex sequencing reactions, in that the polynucleotides to be sequenced are constructed to contain terminal base-specific labeling. Furthermore, the method obviates the need to cross-link reaction products to membranes with chemical reagents. Method steps comprise, parallel separation of sequencing templates or sequencing reaction products from mixture prior to gel electrophoresis, and parallel detection of different samples or reaction products. To achieve separation, plural samples are applied to specific separation channels by parallel sample-specific binding means to said specific channels, thereby combining both multiplexation and parallelism. Detection is achieved by detector means that comprises tag-probe arrays. Suitable devices for performing the method are also described.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

"METHODS AND DEVICES FOR PARALLEL MULTIPLEX POLYNUCLEOTIDE SEQUENCING".

Field of the Invention:

The invention relates to the field of molecular biology, and more specifically the determination of polynucleotide sequence composition.

Related Art:

Due to the importance of the structure of genetic material to the understanding of biological mechanisms and genetic engineering, much effort has been focused on the development of techniques for the determination of the nucleotide base sequence of polynucleic acids, especially of DNA. At present there are only two well established basic methods for the determination of polynucleotide sequence. These are Maxam and Gilbert base specific chemical cleavage[1] and Sanger enzymatic chain termination[2]. Both of these methods serve to generate, from a sample of identical polynucleotide molecules, sub-populations of polynucleotide molecules with a particular one of the four naturally occurring nucleoside bases, according to treatment of the sample aliquot, at the 3′ terminal position. With both of these methods, each such differently terminated sample aliquot comprises a so-called "nested set" of populations of molecules of different length, up to the length of the input sample molecules, but identically having the specified 3′ terminal base composition. Thus, the length of each subpopulation within each such sample aliquot corresponds to a position of the base moiety, for which that sable aliquot has been treated, within the sample sequence. The nested

sets produced by both of these methods are conventionally separated according to length by electrophoresis through gel matrices, generally of polyacrylamide composition. Generally, the sample is labeled with radiolabels or dye moieties such that appropriate detection techniques may localize the position of the bands which are produced by electrophoresis of different length molecules in the gel. By comparison of the pattern generated by such an electrophoresis step, the base sequence of the original sample is inferred. These methods are well known within the field of molecular biology.[3] To improve the rate at which sequence information is accumulated, many different modifications of the original methods described above have been attempted or implemented. These include substituting different resolution media for polyacrylamide gels in the separation step[4], temporal or temporal and spatial separation of fragments involving their departure from the separatory gel matrix (e.g. onto moving membranes or passed detectors), automation of fluid handling, reaction and electrophoresis steps with either conventional slab gels or gels confined to capillary tubes. Despite the improvements in efficiency attained by such modifications, established sequencing technologies are a rate limiting step in much molecular biological research and the several genome projects, including the Human Genome Project. In part, such limitations stem from the characteristics of electrophoretic separation, specifically the time involved and the decreasing resolution power for successively longer fragment length, which decreases approximately logarithmically. It is thus not generally feasible to sequence stretches of longer than 500 bases in length by gel electrophoresis, and much effort and care is required to extend this length beyond 1,000 bases at present.

Multiplex Sequencing:

One innovation which attempts to negotiate the limitations associated with gel electrophoresis is multiplex sequencing.[5] This method has been taught by G.M. Church in U.S. Patent Number 4,942,124 and further by G.M. Church and S. Kieffer-Higgins in U.S. Patent Number 5,149,625. These workers have noted that multiple samples may be separated within the same gel region (lane). When these are uniquely labeled with a tag sequence or otherwise uniquely detectable tag, the length-fractionated sequencing reaction products are then

serially and uniquely detected according to tag identity. Thus, the pattern resulting from the electrophoretic fractionation of the sequencing products from each sample species is accessed by probing for the respective tag. Generally, this method is implemented by transferring or blotting the gel after electrophoresis onto a nitrocellulose or nylon membrane such that the size-separated polynucleotide molecules are deposited onto said membrane with preservation of spatial configuration of the resultant banding pattern. Usually a further step, such as UV illumination of said membrane, is performed to further strengthen the immobilization of said polynucleotide molecules. Probing is generally performed with a radiolabeled, dye labeled or enzyme linked oligonucleotide complementary to the respective tag sequence. After image capture of the banding pattern thus revealed, the oligo probe is denatured from its cognate tag sequence and washed away from the membrane. The membrane is then serially reprobed in the same way until many or all distinctly tagged samples have been detected. Data from each probed sample is used to reconstruct sequence information in the same manner as with Maxam and Gilbert or Sanger sequencing methods. Beyond the reduction of electrophoresis time and effort per sample by this method, distinction with unique tags permits several samples to be pooled together, which enables several samples to be processed in only the four reactions specific to each base, significantly reducing handling steps. A drawback of this method has been the difficulty of automating the serial probing of membranes, though efforts have been made in this regard.[6,7,8]

Electrophoresis methods have been described in a volume entitled *Gel Electrophoresis of Nucleic Acids: A Practical Approach.*[9] These and other methods are taught by J. Sambrook, et al.[10]

Manifold Sample Transfer:

So-called manifold sequencing, a method developed by A. Lagerkvist et al.,[11] reduces manual manipulations in the sequencing of many samples by using modified gel combs or similar structures for the transfer of several samples from sample wells of, for example, microtitre plates, wherein reactions such as PCR and enzymatic sequencing are performed, as w ll as to a slab-gel. Such a method permits handling of more samples than parallel pipettors commercially

available with greater simplicity and economy. This is accomplished through the immobilization of sample molecules to the solid phase of the comb teeth, similarly by a method comprising the step of dipping these into wells initially containing sample material. This multiplies the number of samples which may be separately handled in a given manipulation step by the number of comb-teeth used, and has the further advantage of reducing handling errors, an issue critical in clinical applications of sequencing methods.

Parallel Capillary Electrophoresis:

E.S. Yeung and J.A. Taylor have taught, in U.S. Patent Number 5,324,401, a method for the facile simultaneous use of plural capillary tubes for either capillary gel electrophoresis or capillary zonal electrophoresis. The parallelization of sample separation possible due to the macroscopically small lateral size of these capillaries, as well as the more rapid separations possible by the use of such capillaries and other advantages over conventional slab-gel electrophoresis, facilitate an increase in the number of separations that can be preformed in a given time with given separation effort. A drawback of this method is the lack of correspondingly rapid sample loading methods and means, which limits the ability to take advantage of the parallelism thus enabled.[12]

Nanochannel Array Glass:

R.J. Tonucci et al.[13] have demonstrated a method whereby two or more types of glass, at least one of which is acid-etchable and provided initially in the form of a rod, and one of which is inert, are heated, drawn and fused together repetitively, followed by an annealing step. The number of times such a set of steps is repeated, and the number of contiguous lengths fused together in each step, determine the number of channels ultimately formed, as well as the size of the resulting channels. Channels are formed in a final acid etch step with regularity across the entire medium. These workers note that the definition of the channels formed after etching results from interdiffusion of the different glass types during drawing and annealing, Though these authors do not note this, decreases in such definition entail increases in internal surface roughness of the capillaries thus formed. Note that such articles with channels of

larger than nanometer dimension may be termed a channel array glass most generally, and with channels of micron-range dimension may be termed microchannel array glass.

Microfabricated Separatory Devices:

In recent years there has been increasing interest and effort directed at the microfabrication of analytical instruments and separatory devices. A.T. Woolley and R.A. Mathies[14] have recently described microfabricated capillaries integrated on the surface of a solid substrate. These are formed by patterning and then etching plural trenches in a glass substrate according to photolithographic methods, and then sealing these to form a capillary array by fusion of said substrate with a juxtaposed glass plate. Such microfabricated capillary arrays are then filled, by injection, with a separation medium, for which these workers choose hydroxyethyl cellulose. Using devices with 30-120 micron capillary size, these workers demonstrated the separation of the DNA fragments resulting from HaeIII restriction digestion of phiX174 replicative form DNA in 120 seconds. Capillaries of similar internal dimension but fabricated by conventional means have been used by these workers to separate sequence reaction products[15,16] rapidly over sufficiently broad size ranges with single base resolution, indicating that sequence determination will be feasible with such microfabricated capillary arrays.

Replicated Relief Structures and Microcontact Printing Therewith:

This section reviews related art methods for fabricating structures and decorated articles useful as means in various aspects of the present invention.

<u>Microfabrication with Replication:</u>

Integrated optical devices and other microscale optical components have been fabricated by the replication of relief patterns into polymeric materials[17]. The optical properties of these materials and the structures resulting from pattern replication determine device function and characteristics. Replication methods used include injection molding and casting into polymeric, elastomeric or metallic

molds or mold inserts, and hot embossing with reliefs, for example, metal reliefs. Multiple cycles of replication may be performed to yield a large number of reliefs for the rapid mass production of devices, according to the fidelity of replication and acceptable device tolerances.

Over the last decade, methods have been developed to form three dimensional patterns on the micron and sub-micron scales, combining lithography with electroforming, micromolding and mold replication. An original microrelief is formed by the lithographic depth patterning of a relief material, which is then developed to form the desired predetermined structure (first relief). Electroforming is then performed to yield a metal negative mold insert conforming to the polymer resist structure, and then separated from the original resist. This first generation mold may then be as a mold for the casting or microinjection molding of a second polymeric relief nearly identical to the first relief. The molding or casting process may be repeated with the first generation mold, and the electroforming process may be repeated with second or subsequent generation polymeric relief structures produced with said molds by appropriate methods. Thus, polymeric objects and metallic objects with sub-micron patterns may be economically reproduced. This set of methods, known as LIGA, was developed by E.W. Becker et al.,.[18] These workers reported reproduction of lateral features smaller than 0.1 micron. These methods are particularly useful in the production of devices with microscale features and moving parts. Where the lithographic method employed is synchrotron X-Ray lithography, high aspect ratios may be achieved with resists of up to millimeter depth. More recently, M. Abraham et al.[19] have reported the fabrication of micro-optical systems with LIGA, and further extend these methods by using mold inserts thus produced for hot embossing.

Similarly, M.T. Gale et al.,[20] extend methods from the embossed diffractive foil and compact disk (CD) industries by the replication of an original (e.g. microfabricated) microrelief on the surface of a replication shim through electroforming, followed by replication of this first shim by surface passivation followed again by electroforming. These workers find that the shim replication process occurs with only slight (<2nm) increases in surface roughness per generation, thus enabling the reproduction of nanoscale relief

features, with corresponding cost advantages. Related injection molding methods are described by A. Neyer et al.,[21] and R. Klein and A. Neyer.[22]

Contact Printing of Resist by Lithographic Plate:

In U.S. Patent Number 5,380,620, T. Suzuki and F. Shinozaki teach a method whereby a lithographic plate comprising regions of ink binding and ink repelling regions is contacted with an ink or resist sheet and then contacted with a substrate to which said ink or resist which bound to said ink binding regions of said ink. This method forms lithographic plates by exposing materials similar or identical to those used as resists in conventional microfabrication, such that exposed regions have affinity or repel ink or other liquids, such that a pattern of differential retention of ink or said other liquids may be used to form a pattern of said ink or other liquids on said lithographic plates and then transfer this pattern to the substrate.

Photolithography Generated Relief Patterns, Mechanical Transfer Thereof, and Use as Lithographic Plate:

Methods for the production and transfer of relief patterns in polymeric resist materials or variants thereof are reviewed by B. Bednar, J. Kralicek and J. Zachoval.[23] A relief is produced by selectively exposing regions of a photoresist coated onto a first surface and mechanically transferring either the exposed or non-exposed regions to a polymeric foil second surface by juxtaposing said second surface to the image-exposed polymeric resist and relying on the differential adhesion properties resulting from exposure. Thus, a positive image is produced on one surface while the other surface retains the corresponding negative pattern. Such relief patterns may provide differential wetting or liquid-retention properties differing from those of the underlying surfaces, and thus be used as offset lithographic plates.

Microcontact Printing with Elastomeric Microreliefs:

G.M. Whitesides and co-workers have done considerable work developing methods for the use of replicated elastomeric microreliefs for microcontact printing.[24] It is presently important to note that

this work demonstrates the conformation of the raised portions of an elastomeric microrelief surface to the shape of another surface to which it is juxtaposed and impressed. These workers have similarly used such reliefs for micromolding, which further indicates that conformed contacts resulting from impression of such microreliefs onto a juxtaposed surface forms a liquid-tight seal.

Chemical Manipulation and Synthesis with Scanning Probe Microscopes:

W.T. Müller, P.G. Schultz et al.[25] describe a patterning process whereby a surface is adsorbed with molecules comprising a chemical functional group which may be transformed to a second chemical functionality by contact with a metal catalyst, an AFM is used to scan a tip coated with said metal catalyst over areas of said surface which are intended to be transformed, and said chemical functional groups transformed by the resulting, spatially limited catalytic process are then subjected to further chemical reactions which are selective for the catalytic reaction product second chemical functional groups. In particular, these workers situate azide (alkylazide) functionalities on a substrate surface, and contact a platinum coated AFM tip to these in the presence of dissolved H_2, to effect the catalytic hydrogenation of those of said nitriles contacted by said AFM tip under positional control effected by an AFM. The nitrile groups thus contacted are reduced to amines, which these workers detect by reaction to aldehyde modified fluorescently labeled beads, which are observed to bind only to regions which underwent said contact with said platinum. B.J. McIntyre, M. Salmeron and G.A. Somorjai[26] have similarly effected positional control of metallic surface catalysis using a platinum-rhodium STM tip in an atmospheric-pressure chemical reactor to catalyze the rehydrogenation of molecules comprising alkyne functionalities on (111) platinum surfaces.

Alternative Sequencing Methods:

Various novel methods for the determination of polynucleotide sequence have been proposed. These include substitution of mass spectroscopy[27] (including time-of-flight mass spectroscopy) separation for gel electrophoretic separation, exonucleolytic degradation with

nucleotide transport and single molecule detection of the effluent[28], scanning probe microscopic visualization of polynucleotides at high resolution[29,30] with the aim of base discrimination within a single molecule, and various sequencing by hybridization methodologies. Of these, only the last appears at present to have significant practical potential, but significant obstacles and challenges remain.

Sequencing by Hybridization:

Two general variants of sequencing by hybridization methodology have been proposed. In the first, a sample under investigation is hybridized to an array of oligonucleotides situated on a surface, fabricated, for example, by the spatially controlled light directed synthesis methods of S.P.A. Fodor et al.,[31], where the composition of an array element is related to the position of said element within said array of oligonucleotides. For example, an array may comprise all possible 10-mer oligos, in which case there are over 10^6 array elements. An unknown sample is denatured and permitted to anneal with said array under stringent conditions. Said array is examined for the presence of bound sample molecules, generally by optical means. Ideally, the sequence composition of a sufficiently short sample molecule may be reconstructed by using information about which array elements have bound sample molecules and about the compositions of each of these array elements, with algorithms which overlap each of the short sequences (corresponding to the 10-mer oligos in this example) to produce a linear sequence. Two difficulties arise with this method. First, because DNA includes tandem and dispersed repeat elements (including transposons), a particular 10-mer, which is one of the more than 10^6 possible 10-mers may occur many times in naturally occurring sequences much shorter than a megabase.[32] Thus, this method will lead to branch points which ambiguate the sequence data thus obtained. Longer oligo probe libraries or arrays will contain more elements, which thus occur less frequently, but require more surface area and synthesis steps, and will further diminish the ability to discriminate between perfect and mismatched hybridization. Second, even for shorter oligos, such as 7-mers, hybridization is not perfect for all sequences under any given conditions, as is clear to those familiar with other probe hybridization techniques such as northern blot gel analysis; optimal conditions for a particular probe

to form perfect hybrids and not form imperfect or mismatched hybrids vary in a sequence dependent manner, and it is not always simple to predict this optimum even where only a single probe is concerned. Thus, for any given array of all possible n-mers, only a subset of elements will display optimal binding discrimination characteristics under any set of particular conditions. Further, because of the formation of imperfect mismatches, discrimination between strongly and weekly binding probes must be considered rather than simply the success or failure of an array element to bind sample molecules.[33] Various attempts have been made to address these issues, with some success[34], but the technique is not yet competitive with established methods.

In the second variant of sequencing by hybridization methodology, the sample sequence (which is generally in the form of a phage or plasmid library) is immobilized, generally in the form of bacteriophage plaque or bacterial colony replicas, on a membrane, which is serially probed with a different labeled oligonucleotide of known sequence and length n. Here, because only one kind of hybrid is formed during each step, conditions may be closely optimized for homoduplex formation in each probing step. Plaques or colonies forming hybrids are noted. Repetition over the full combinatorial library of oligonucleotides of sufficient complexity relative to the complexity of the sample thus provides information about all of the n-mer sequences present in a particular plaque or colony. After each element in the full combinatorial library has been used to probe the sample sequence, linear sequence information may be reconstructed as in the first variant of sequencing by hybridization methodology. The two main drawbacks of this method are the requirement for synthesis of large probe libraries, which is much more laborious than the production of surface immobilized arrays by optically patterned deprotection[35], and that significant linear sequence information may only be reassembled only after many (thousands or more) probing steps.

The first variant of sequencing by hybridization is, however, proposed as a confirmatory sequencing method to identify any errors in sequence information obtained by other methods, and the same methodology is proposed as a means of classifying polynucleotides in unknown samples, for example, to detect the presence and type of

polynucleotides present in clinical isolates. In this latter application, even where a particular polynucleotide displays a hybridization pattern which permits only ambiguous sequence reconstruction, a characteristic pattern which may serve as a highly specific signature may nonetheless be observed.

Parallel Multiplex Polynucleotide Sequencing:

In U.S. Patent Application 08/445,094 filed on 19 May 1995, E.M. Rabani teaches a method for the sequencing of polynucleotides, especially DNA, combining the separation throughput advantages gained in multiplex methods with the detection methods made possible by oligonucleotide arrays. In the most general aspect of that invention, separation is performed so as to effect length dependent fractionation of complex mixtures of tag-primer primed sequencing reaction products, which said fractions are separately scored for the presence of tags, or for labeling moieties in communication with tags, on oligonucleotide arrays which comprise tag-probe sequences (i.e. such arrays may comprise as many distinct tag-probe oligonucleotides as there are distinct elements in such arrays, which may exceed 10^6 or 10^8 different oligonucleotide elements per cm^2, depending on the fabrication methods used to produce these.) In a preferred embodiment, said tag-primer primed sequencing reaction products are affinity labeled during that reaction with one of four distinct labeling moieties the identity of which corresponds to the identity of the terminal base moiety. After fraction collection, hybridization to an array comprising said probe-tags, and washing steps, four differently colored fluorescent beads each distinctly derivatized with one of four corresponding receptor types, which bind one of said distinct labeling moieties, are contacted to said array, with the result that array elements to which said tag-primer primed sequencing reaction products have bound are labeled, in this instance by a fluorescent bead. The color or identity of the bead bound to a particular array element or tag-probe reveals the identity of the terminal base moiety of the sequencing reaction product fragment thereupon, and hence the identity of the base at the position within the sequence of the tag-primer primed sequencing fragment probed by said tag probe (corresponding to said particular array element) corresponding to the position of said fraction within the series of

fractions collected (in a manner precluding the presence of successive base terminated species of the same tag-primer primed sequencing reaction product in the same fraction) and more particularly within the ordered set of labeling data obtained from similar scoring of all successive fractions. With this method and the several variations disclosed therein, multiplexation advantages in sample preparation, sequencing reaction and separation are combined with the rate advantages of parallel detection, such as is possible with said large (i.e. complex) probe-tag arrays and the technical facility of photolabel detection (e.g. fluorescence detection, especially of fluorescent beads, with appropriate CCD arrays with data capture by an associated digital computer and storage device.) These advantages come primarily at the cost of the requirement to examine a large number of fractions (greater than or equal in number to the length in bases of the linear sequence to be obtained for each or most tag-primer primed samples.) At the array hybridization step, fractions may be applied successively to a single array, which may be referred to as serial fraction array tag-probing, or multiple fractions may be applied (on a one to one basis, preserving information regarding ordering) to plural said tag-probe arrays, which may be referred to as parallel fraction array tag-probing. Additionally, data regarding labeling of array elements (in association with ordering according to fraction) may be collected from such tag-probe arrays sequentially by a single suitable photodetection means or in parallel by multiple such photodetection means. CCD based detection is, nonetheless inherently parallel, so this latter distinction refers to hierarchy of detection parallelism. Thus, despite the great advantages in overall sequencing rate attainable by such methods over conventional and other alternative methods, the sequence data collection rate with this method remains proportional to the complexity of the probe arrays used (i.e. the number of distinct probes per array), but the effort or number of automated steps required is primarily determined by the number of fractions thus examined. It would be desirable to extend this method using appropriate means so as to reduce or eliminate the need for or the number of separate fractions which must be separat ly handled (whether manually or by automated means) without thus reducing the rate at which data are collected or the economy of sequence

determination by such a method. It may for some purposes be further desirable to rely on probe arrays of generally smaller diversity (i.e. consisting of fewer array elements), since the effort required to fabricate these is primarily proportional to the length of the variable portions of the oligonucleotides therein, and further since more methods are capable of efficiently producing lower density arrays.

Analysis of Conventional Methods:
Logical Steps of Nested Fragment Set-Based Sequencing:

This discussion delineates the essence of conventional sequencing methods in a manner which will facilitate the understanding of the present invention both in its full generality and the large number of variations comprehended by the present invention. In both enzymatic and chemical methods, samples are in a first step prepared as suitable reactant molecules (e.g. as single stranded and primed molecules for enzymatic methods or as end-labeled fragments for chemical methods). In a second step, these samples are subjected to desired sets of reactions or combinations thereof to form fragments which are either separate nested sets each comprising only fragments with a single terminal base moiety or fragments differently labeled according to the identity of the terminal base moiety which said fragments differently labeled comprise. In a third step, these samples are each applied to a different separation channel. In a fourth step, samples are separated into the constituent molecular species which the nested set comprises according to length or mass separation (e.g. electrophoretic mobility or transport). In a fifth step the distinct said constituent molecular species in each separation channel are detected (either according to the final position reached in said separation channel at the time said fourth step was concluded by spatial detection method, the coordinate of which corresponds to the order of molecular species along the length of said sequencing channel, or according to the time or time delay at which a fragment is transported passed a detector situated at a fixed position along the length of a separation channel, concurrently with the separation step, in which case the separatory coordinate corresponds to the temporal ordering of transport dependent events;

in short, these are spatial and temporal separation coordinate resolution methods, respectively.) As more broadly considered, a sixth step consists in the assembly of sequence information obtained in the preceding steps into longer stretches of sequence data where the sequences of distinct samples presents both sufficient identity and unambiguous or unique match at the putative sites of overlap. In summary, these steps are:

1. sample preparation, which may include creation of libraries or plasmids, growth, purification and/or amplification steps;

2. base specific fragment generation (sequencing reaction), which may be effected by enzymatic or chemical methods;

3. application to separation channel;

4. separation, which may avail slab gel electrophoresis, capillary tube electrophoresis, microfabricated channel electrophoresis, or mass spectroscopic separation;

5. detection of separated species; and,

6. reconstruction of sequence information from separation pattern information.

Automated implementation of some or all of these steps have increased efficiency and reduced labor requirements and costs per base of sequence information, generally only modestly departing from the above basic scheme. Capillary array electrophoresis yields gains in parallelism at step 4 and thereafter. The original multiplex sequencing protocol of Church and Gilbert, and known art variations thereupon may gain advantages in reduction of sample handling and separation starting at step 1 above, gain these advantages at the cost of repetitions of serial reprobing steps in step 5. Multiplex PCR may reduce the number of samples which must distinctly be handled in step 1. Manifold transfers may reduce the number of manipulations as the results of step 1 are subjected to step 2, but still require the same number of vessels and separatory channels. The various encoding schemes (described below) which have been used to permit the sequencing products of two to four sequencing reactions yielding fragments terminating at a predetermined base moieties in predetermined proportion and/or with distinct, base-specific or reaction specific labels to be separated in the same separation channel, up to four, while n t confounding data collection with

appropriate detection means, have provided two- to four-fold more efficient utilization of however many separation channels are thus used, primarily gaining advantage in step 4, but also in step 5. CCD arrays, photon cameras or other suitable electronic devices for image capture having increased resolution may improve the detection rates in protocols using optical labels including fluorescent labels. Improved image recognition algorithms and faster computers (including parallel computers) may increase rate of the final step of reconstruction of data from individual samples to ordered data about the original sample sequence.

Variations providing sequencing rate, economy advantages or reduction of manual steps:

Each of the aforementioned seven steps, as well as other associated steps and methods, admits either multiplexing or parallelization or both, with different advantages and tradeoffs according to the sequencing task intended. In conventional practice, slab gel sequencing admits the small-scale parallelism of multiple separation channels (i.e. lanes). As will be clear from the foregoing description, the multiplex sequencing methods of Church and Gilbert may avail multiplexing in sample preparation and reaction and separation, or reaction and separation, or only during separation; detection is performed serially, i.e. generally one sample per lane per detection step, where serial reprobing is made possible by the association of a unique tag sequence, generally as part of the priming molecule, though in the original method part of the sample sequence, by a labeled probe having a complementary sequence to said tag sequence, which such labeled probe will be referred to herein as a tag-probe. At the levels of separation detection, methods such as the encoding of terminator concentration or relative activity with corresponding effects on band intensity (e.g. a binary encoding scheme such as 2A+G and G+2C reactions of the same primer-template end or length labeled identically and run in adjacent lane such that 0,0=T; 2,0=A; 0,2=C; 1,1=G) permit two-fold advantages in the form of separation multiplexation (though here in the limited sense attaching to fragments derived from a single primer-template complex or the like) and corresponding advantages in fragment detection rate with sufficient intensity discrimination. More recently, the use of dye

molecule labels or reporters, which may be in communication with primers used (end-labeling), or with deoxynucleotides (length-labeling) permit the separation of the combined mixture of products of four separate sequencing reactions (each corresponding to a different chain terminator moiety and having been labeled with a dye moiety having a distinct absorption and/or emission spectrum) in a single lane, permitting further improvements in this limited sort of multiplexation of separation and detection. The extent of such multiplexation based on distinct labeling is limited by the number of suitable spectrally distinct dyes available. This category of base-moiety multiplexation limited to an individual primer-template sample has, however, been extended, through the use of distinct fluorescent or other optical labels in communication with each of the four chain terminator analog (generally, 2'-3'-dideoxy nucleotide triphosphates) such that the same synthesis event which terminates chain elongation also places a label corresponding to the identity of the incorporated chain terminator in communication with the resulting terminated elongation fragment. This further improvement results in a four-fold reduction of the number of reactions which must be performed to sequence a template sample. Even in this last case, such methods, by themselves, only provide four-fold advantages through the steps of sequencing reactions, separation steps and detection, providing only four-fold advantage at the system throughput level.

Object of the Invention:

It is an object of the present invention to gain the advantage of parallelism in the detection steps of multiplex sequencing and thus better exploit the parallelism of sample handling gained in sample preparation and size fractionation steps with multiplex sequencing.

It is a further object of the present invention to provide highly parallel base position detection without extensive use of high-complexity oligonucleotide arrays, for example, those comprising over 10^6 elements each, so as to provide the possibility of a comparable rate of detection with a larger number of oligonucleotide arrays each consisting of fewer elements.

It is a further object of th present invention to provide for the multiplexation of sample loading in parallel separation methods including parallel capillary electrophoresis methods.

It is a further object of the present invention to combine the parallelism advantages of parallel capillary electrophoresis with th sample processing advantages of multiplex separation.

It is a further object of the present invention to provide means to combine the advantages of parallel capillary electrophoresis and multiplex separation with the advantages of parallel detection. In particular, it is a further object of the present invention to minimize or eliminate the need to perform multiple manipulations to form separate liquid fractions for analysis on separate probe arrays while still gaining the parallelism advantages of detection comprising the use of probe array scoring of plural distinct species occurring within a size fraction.

It is a further object of the present invention to provide rapid and inexpensive methods and means by which complex polynucleotide samples including large portions of genomes may be analyzed with single base resolution.

It is a yet further object of the present invention to provide rapid and inexpensive methods and means for the high resolution analysis of many independent samples of genetic material, for example, one or more particular loci from many individuals within a population of organisms.

Methodologies and means fulfilling one or more of these objectives would increase the rate of progress of much biological research and applications including recombinant DNA technology, and have further relevance to both basic medical research and clinical medicine.

Summary of the Invention:

The essence of the present invention is to combine any of the various multiplexing methods possible at each necessary step of Sanger enzymatic sequencing protocols or Maxam and Gilbert chemical sequencing protocols for use with appropriately prepared complex samples, with the advantageous degrees of parallelism at appropriate steps of sequencing protocols. Prominent multiplex methods among these are multiplexing of samples in reactions directed at generating terminal base specific nested fragment sets or terminal base specific reporter-linked fragment sets, labeling and detection, as well as use of variations on aspects of the multiplex sequencing methods of Church and Gilbert obviating the crosslinking of sample sequencing

reaction products to membranes with chemical reagents. Prominent parallelization strategies enabled include parallel separation of sequencing templates or sequencing reaction products from mixtures prior to molecular mass fractionation (e.g. slab or capillary gel electrophoresis, capillary zonal electrophoresis, mass spectroscopy), parallel separation steps through the use of plural separation channels or means, and parallel detection of different samples or reaction products generated from different samples within the same mixture or mixture fraction. At the level of separation, single or plural sample mixtures or sample reaction-product mixtures may be prepared in multiplex, may each be separated in a manner directing them to specific separation channels or other separation means by parallel sample-specific binding means and thereafter separated by plural said specific separation channels or other separation means, combining both multiplexation and parallelism. At the level of detection, separation channels may comprise or suitably articulated with sample recovery means in turn comprising parallel sample separation means, may be detected by parallel detector means (e.g. tag-probe oligonucleotide arrays and photodetector arrays such as CCDs for visible labels) and alternatively or further discriminated by signal amplitude encoding schemes where appropriate with detection methods used. Suitable devices and apparata and variations thereupon facilitating such methods are further described.

Description of the Invention:

Definitions:

As used herein, the term parallelism will generally indicate concurrent performance of similar operations on different populations or sub-populations of molecules or samples across space. Parallelism thus entails the multiplication of devices or articles with the methods of the present invention to gain a corresponding multiplication in throughput.

In contrast, as used herein, the term multiplex will generally indicate the concurrent manipulation of different samples (or molecular products yielded by reactions performed collectively upon the ensemble of said different samples) within a continuous region of space, area or fluid volume. Multiplexing thus involves the

reduction of handling steps, and the ratio of devices or articles used to accomplish a multiplexed step per sample, by treating different samples identically, which is possible where each of said different samples comprise unique features facilitating subsequent separation and/or discrimination according to the origins of molecules (or of molecular products yielded by reactions performed collectively upon the ensemble of said different samples) of from each of said different samples. This is stressed because within much of the related art the term "multiplex" has been applied to arrangements or methods which are more accurately described as parallel. Note that the present usage is consistent with that employed by Church and Gilbert.

Thus, it is important to note that in some regards, parallelism may be exploited as a method of de-multiplexing differently derived fractions of sample mixtures or the like as well as demultiplexing the information obtained during detection steps.

These distinctions will facilitate a clearer understanding of the essence of the present invention as it is set forth below and as it may be variously embodied by those of ordinary skill in the relevant fields of related art (which include molecular biology, recombinant DNA technology, microfabrication, integrated optics replication, combinatorial chemistry, systems operations theory, electronics and electronic sensing, and, computer architecture engineering.)

Further note that the term separation channel is used generically to include such diverse separations means as a mass spectrometer, a slab gel lane, a capillary gel, a capillary or the like. By contrast, the terms formed channel, formed relief channel and articulated channel, all of which may serve as separation channels and may be of varied form and composition, do not apply to structures formed from conventional capillary tubes. Lateral closure of such channels need not be permanent, and thus separated species may be recovered by laterally accessing said thus separated species by reversing said lateral closure of such channels, i.e. separating one or more portions of materials making up said channels so as to expose a substantial proportion of the internal volume of said channels. Conventional capillary tubes, with which conventional capillary gel electrophoresis and capillary zonal electrophoresis are accomplished,

in contrast have only terminal openings and generally poorly accessible internal surfaces.

Brief Description of the Figures:

Figure 1 depicts demultiplexation array means (10). In Figure 1a, said demultiplexation means are shown in a sealed channel formed by juxtaposition of trench (20), while in Figure 1b, said demultiplexation means are shown each in one among many plural sealed channels formed by juxtaposition of trenches (30) which thus form a separatory channel array. Figure 1c shows an isometric cross-section of one of said separatory channels depicted in Figure 1b.

Figure 2 depicts a bin array (100) comprising a grounded electrode contact (110) across the bottom surface of said bin array (shown with schematic symbol) and shows a portion of a transparent channel array (140) of appropriate pitch offset from said bin array.

Figure 3 depicts a superimposed view of and array (two sets of grooves shown) comprising fraction transfer grooves and plural surface immobilized tag-probes (330) of a striped-patterned tag-probe array, and electrodes (350a and 350b) for lateral electroelution, positioned facingly over the trenched surface for forming a separatory channel array, with one such trench shown (320). Note that features 400 and 420 are extensions of the surface comprising said fraction transfer grooves which serve to define two distinct openings to said separation channel trenches when the fraction transfer grove array surface is facingly juxtaposed to the separatory channel trench array surface, and further (420) to prevent cross-transfer of fractions.

General Description of the Invention:

As recited above, established methods for the determination of sequence composition of DNA or other polynucleotide samples generally comprise six categories of steps:

1. sample preparation;
2. base specific fragment generation (sequencing reaction);
3. application to separation channel;
4. separation;
5. detecti n of separated species;

6. reconstruction of sequence information from separation pattern information.

Heretofore, with only a few exceptions, overall methods for the determination of sequence composition of DNA or other polynucleotide samples have availed multiplex handling or parallel manipulations at only one or a few of the above categories of steps. This is attributable to a general lack of previously identified means for sample demultiplexation, provided in the present invention, or to a lack of parallel detection of plural species (of similar length in nucleotides) separated simultaneously in the same separation channel, such as those recently described by E.M. Rabani in the above cited patent application. Demultiplexing and parallel detection of distinct species of similar length simultaneously separated in the same separation channel individually, and especially in combination, provide the crucial links necessary to exploit combinations of parallelism and multiplexing, in many possible combinations, in order to multiplicatively increase sequence determination rate and improve sequence determination economies, expressed as cost per finished base.

Thus, for example, samples may be prepared (and as necessary, tagged) in multiplex (including multiplex amplification), subjected to sequencing reactions and/or labeling in multiplex, partially or completely demultiplexed with predetermined transfer to parallel separation channels, separated in multiplex in each of said parallel separation channels or merely in parallel with only one sequence per channel, and detected in parallel (including, as applicable, parallel detection of or by parallel detection of samples separated in multiplex, for each separation channel, favorably performed simultaneously and hence in parallel for all separation channels.)

Manifold Loading of Parallel Separation Means:

This aspect of the present invention combines manifold loading of samples with the use of capillary arrays or other parallel separatory channels or separatory channel arrays. Manifolds may be one dimensional combs as previously described by A. Lagerkvist et al.,[36] or may be two dimensional regular arrays of teeth (e.g. rods) projecting from a two dimensional surface, fabricated by suitable

means such as molding, including LIGA techniques or variations thereupon, or other suitable means.

In a second variation, manifold separation may further provide separation of probe targeted species from complex mixtures. Oligo-probes are situated on manifold teeth and used to fish for probed sequences from a complex pool, e.g. several libraries or complex samples of different origin are individually contacted with one comb tooth, such that a comb may carry several separated sample fragments in a manner ordered according to which tooth was contacted with which sample, which is particularly useful for application to parallel separation channels. This second variation is actually a specific case of demultiplexation, and manifolds with plural distinct oligos situated on said teeth may be seen as a specific case of demultiplexation arrays when used in this way, with multiplexed samples. This variation however suffers from shortcomings associated with not integrating probe-tags with separatory channel arrays, particularly that each of said teeth must be contacted with each sample from which it is to capture sample material, as opposed to the unitary contacting provided by the demultiplexation arrays of the present invention.

<u>Random Loading of Parallel Channels With Parallel Amplification:</u>

The present aspect of the present invention involves the use of one- or two- dimensional arrays of sample bins, wells or chambers (hereafter referred to as bins). A single sample volume, comprising a number of (individual) molecules equal to the number of wells used, is distributed among these bins. Alternatively, a sample volume, containing an appropriate concentration of (distinct) sample molecules, and larger than the total volume of all bins of said array, is applied to the array of bins in a manner such that a predetermined portion of said sample volume sufficiently fills each of said bins of said array, such that on average each bin contains one sample molecule; the excess of said sample volume is drawn off of the surface of said array. These aspects of the present invention may be termed Poisson statistical random sample deposition, or simply random isolation. Samples are then (if they had not previously been, or are not otherwise suitable for amplification with standard

primers) ligated to PCR primers, and then PCR or other amplification is performed in each chamber, where thermal cycling of all chambers is effected simultaneously, followed by reactions to generate nested set sequencing products with appropriate labels. Other means of effecting parallel amplification are, of course, similarly comprehended by the present aspect of the present invention. A sufficient number of PCR amplification cycles are performed to yield a quantity of DNA that will ensure that sufficient signal will result from the labels in each sequencing reaction fragment for reliable detection. Thus, samples comprising distinct sample molecules which each comprise substantially different insert sequences or are each derived from substantially different regions of a genome, are generated in an ensemble of bins or chambers (e.g. in a microtitre plate or higher density chamber array); according to Poisson statistics, approximately one third of said chambers contain sequences amplified from only one molecule. The remaining chambers will contain material amplified for either zero or two or more sample molecules. Said array of bins is designed so as to conveniently be juxtaposed to a separatory channel array, or alternatively to be used with manifold transfer means for transfer to a separatory channel array; i.e. said array of bins is arranged in either a mirror image of the configuration of entry ports to said separatory channel array with identical array pitch (for transfer by juxtaposition), or arranged identically to the configuration and pitch of said separatory channel array. In the case of transfer by juxtaposition, physical transfer may take place by capillary action and fluid flow, effected by suction at the exit ports of said separatory array where said separatory array comprises entry ports which may be immersed in said bins, or physical transfer of sample molecules may take place under the action of an applied electrical field. In the case of transfer with manifold, denaturation may free unlinked strands from the teeth of said manifold, as in the method of Lagerkvist *et al*.,.

It should be noted that the approximately two-thirds inefficiency of random bin filling, in absence of other measures to fill unfilled bins with single samples, may be greatly offset in various embodiments by the ability to load samples simply and in parallel into a very large number of parallel separation channels.

Note that the method and means of the present aspect of the present invention may provide further advantages in cases where a sample is in the form of a complex heterogeneous mixture (e.g. a genomic library or a cDNA library comprising thousands to billions of different chimeric clones) having an estimated number of distinct clones per unit volume. Clones may be isolated by such methods availing statistical separation by first diluting an appropriate volume of said complex heterogeneous mixture such that the desired number of molecules are present in a appropriate volume suitable for application to a bin array according to one of the two variations described above. On average, different clones will appear in a bin array in proportion to their relative copy number or representation within said complex heterogeneous mixture. For example, shotgun libraries may be thus sampled and separated into bins each containing one individual sample molecule, amplified, and directly applied to parallel separation channels or means. Further, transformation steps or phage packing steps may be eliminated in certain cases where a library need not be maintained, by randomly fragmenting genomic material, subjecting to treatment with phosphatase, and ligating into vectors comprising appropriate amplification primer sequences, sequencing primer sequences, and optionally, tag sequences, and then separating into bin arrays for amplification. This eliminates host-induced biases arising in library growth. Thus, genomic material may be shot-gun sequenced directly, without even any transformation steps, providing important rate and cost advantages in certain applications. cDNA libraries may similarly be sequenced without transformation or phage packaging and growth steps. Multiplex PCR products of genomic material may similarly be sequenced directly, for example yielding sequence information regarding several regions of each sample genome.

Various assays (e.g. detection of pyrophosphate, a product of polynucleotide polymerization arising in enzymatic sequencing reactions, and most favorably colorimetric pyrophosphate assays) may be performed on all bins subsequent to sequencing reactions and prior to loading to separation channels to determine whether any polymerization took place within each bin, and further assess the degree of reaction which will depend in part on the number of individual molecules randomly distributed into each of said bins.

Thus, empty bins may be readily detected, and with sufficiently quantitative photodetection methods, bins containing material amplified from two or more individual polynucleotide molecules may similarly be detected. Such photodetection of polymerization in said bin arrays is favorably accomplished with photon cameras or CCDs or the like interfaced with one or more digital computer. With an appropriate arrangement and electrical or electronic design of individually controllable electrodes and electrode controlling means, or of individually controllable parallel fluid transfer means (e.g. comprising controllable valves), parallel transfer of samples from each bin of said bin array to the corresponding separatory channel of parallel separation means may be individually controlled. The individual transfer controlling means of such an arrangement may be interfaced to one or more digital computer or the like, which may use data gathered as described above or equivalent data pertaining to the occupancy of each bin to control sample transfer of the contents of each bin of said bin array to said parallel separatory channels or separatory channel array such that only the contents of bins containing exactly one or alternatively fewer than two distinct samples are transferred to said separation means, while data concerning which separation channels have not thus received samples are further maintained. Such a procedure is repeated with additional bin arrays, such that separatory channels not previously receiving samples will receive samples from any of said additional bin array if and only if the juxtaposed or otherwise corresponding bin contains either exactly one or alternatively fewer than two distinct samples. With each repetition of such a transfer cycle with different bin arrays, the probability that any given separatory channel has not received an individual sample decreases by a factor of approximately one-third. Thus, after n transfer cycles, the probability that any separatory channel of said parallel separation means or separatory channel array has received sample material amplified from exactly one molecules is given by $1-(2/3)^n$, for $n \geq 0$, which indicates that after 4 transfer cycles over 80% of said separatory channels will have received sample material amplified from exactly one molecules, which percentage increases to over 91% after 6 transfer cycles. Because application of samples to bin arrays, amplification of material in bin arrays, and transfer of contents of bins to separatory channels

are not involved steps with appropriate apparatus, such repetitions of bin-array transfer cycles may not represent significant costs of time or effort which maximizing the quantity of useful information obtained per unit time from a parallel separatory channel array of a given degree of parallelism.

Loading of Multiplex Samples onto Parallel Separation Channels With Demultiplexation:

The present aspect of the present invention uses one or more immobilized probes or tag-probes (which are generally oligonucleotides, but may, for example be sequence specific binding proteins, etc.) to selectively cause only polynucleotide sample molecules comprising complementary sequences to bind to a selected surface region or spatial region, to facilitate the transfer of only bound polynucleotide molecules to a particular separation channel. The ensemble of distinct said immobilized probes or tag-probes, each of which is chosen and/or synthesized to be substantially distinct from any other probes or probe-tags of the ensemble, are referred to as a demultiplexation array, which term also refers to the surface and object on which they are situated when used most generally. Generally, said probes or tag-probes are provided as a linear array and are situated with a pitch equal to that of the entry ports of a separatory channel array such each element of said demultiplexation array is in, or may be placed in, convenient proximity to exactly one of said entry ports, and such that upon release of said sample molecules from said probes or tag-probes, said sample molecules are either situated within or are readily transported in a targeted manner to the corresponding said entry port. For the case of oligonucleotide based demultiplexation arrays, release of bound sample molecules may, for example, be effected by increasing temperature sufficiently, or exposure to a sufficient concentration of denaturing reagents (e.g. 7M urea).

A complex sample (termed a multiplexed sample) is contacted as a solution with said demultiplexing array such that the sample molecules therein may bind to complementary probes and such that most or all of the probe elements of said demultiplexation array will bind sample molecules of substantially identical origin. For example a flat surface with a linear trench may be juxtaposed with said trench

oriented lengthwise juxtaposedly along said demultiplexation array such that said trench passes over all array elements and forms a channel. A sample mixture containing sequencing reaction product fragments, for example, which are derived from plural distinct sequence regions (i.e. the products of a multiplex sequencing reaction in this example) is flowed through the channel thus formed such that it passes over each of the elements of said demultiplexation array, and non-denaturing conditions of sufficient stringency are maintained. Then a rinse solution is passed through said channel thus formed to carry away and unbound molecules, as non-denaturing conditions of sufficient stringency are maintained. Said array elements of said demultiplexation arrays are designed and produced so as to comprise a sufficient number of probe molecules to capture a sufficient number of sequencing reaction product fragments for reliable detection of each species corresponding to termination at a particular base with the detection method used. Said sample mixture is prepared so as to either not have any two identical tag sequences in communication with different sample sequences to eliminate probing ambiguity, or is produced such that probing ambiguity will be improbable. Alternatively, said probe sequences or other binding means are chosen so as to target only one sequence in said multiplex sample based either on knowledge about the sample or statistical predictions concerning probe sequence occurrence. Said flat surface with a linear trench is then removed from said demultiplexation array, which is then positioned juxtaposedly to the array of entry ports of a separatory channel array, if not part of the same structure. If said demultiplexation array is situated on the same surface of an article comprising separatory channels as the input ports of said channels, then a surface conforming to said demultiplexation array and input port array surface but comprising trenches for juxtaposedly forming channels permitting the transport of molecules from each of said demultiplexation array elements to the corresponding separatory channel input port is then appropriately juxtaposed, and either temperature is briefly increased or denaturants are introduced into said trenches, in either case as the electrophoresis field is applied to initiate separation of the thus demultiplexed sequencing reaction products in parallel.

Distinct tag-probe oligos or internal sequence probes or other such probes of different composition are situated at the entry port to each of two or more, and preferably several, and most preferably thousands or more, separation channels which may be capillary tubes or may be a channel array glass (or microchannel or nanochannel array glass) or other types of separation channels.

In the case of these channel array glass types, probe-tag arrays, which may favorably be of triangular or hexagonal geometry, are either fabricated on the surface of said channel array glass, which surface comprises said entry ports or openings, or are most favorably situated on a surface which is juxtaposed to the surface comprising openings of said channel array glass in a loading step.

Complex mixtures of different sample molecules are first bound to said probe tag arrays or demultiplexation arrays, and unbound species are removed by one or more wash steps. In the case of a probe-tag array which is situated on an article or surface other than that of said channel glass array, sample molecules are bound to said probe tag array which is then washed before juxtaposition or the contacting of said probe-tag array with said channel array glass. An electric field directed along the axis of said separatory channels and oriented with an increasingly positive electrical potential from said probe-array to the distal opening (exit port) of said separatory channels is established, and conditions (e.g. temperature) are adjusted to disrupt the binding of said sample molecules with said probe-tag array such that sample molecules are freed and immediately undergo field directed transport into and along the said separatory channels situated closest to the respective probe-tag array element. Thus, sample molecules enter particular separatory channels according to the array element to which said sample molecules first bound, and according to which separatory channel said array element was juxtaposed. Thus, means, including automated instrumentation, embodying this aspect of the present invention should preferably juxtapose such arrays in a repeatable manner to said surface such that the sample molecules bound to a known array element will be separated by a particular, thus predetermined, separatory channel.

Here, the problem of loading a particular sample (or set of samples in the case of multiplex separations) into one particular separatory channel out of many is solved by situating a binding

activity in proximity to the entry port of a separatory channel. A variation comprises the transfer of sample molecules from a probe-tag array to a surface of a manifold, with binding of sample molecules to said surface of a manifold, after which said manifold is used to carry sample molecules into arrays of vessels, if any, to perform reactions, if any, and comprises transfer of sample molecules or reaction products therefrom to separatory channel arrays, channel glass arrays or capillary arrays.

Using such arrays, two sample fraction detection schemes are possible. The first involves photodetection of molecular species, which may involve species labeled with fluorescent dyes. In this case said channel array glass must be transparent at least to the excitation frequencies used, and where emitted photons traveling in any direction other than that of the channel axis direction, further transparent to the wavelength of emission of the dyes used. Fractionated species may be detected by scanning a laser beam of appropriate frequency through said arrays, and detecting fluorescence of fluorescent labels as a function of beam position and linear position along said beam, as proposed by S. Esener et al.[37] for three dimensional optical memories with one-beam addressing. Such a one-beam approach may comprise the scanning of said one beam across planes comprising the axis of each of many channels or across a plurality of channels at a fixed position along the length of all such channels as separation proceeds, and in either case recording the dye emission signals thus associated with particular positions or times with each channel of said array. Planar beams, also used by Esener et al. for two dimensional readout, may similarly be used with two-dimensional photodetection means (e.g. CCD array camera) to either collect spatial separation data along the lengths of those channels whose axes substantially occur in the plane of said planar beams, or to collect signals regarding the passage of species through said plane (i.e. according to a temporal separation coordinate), here oriented perpendicularly to the axis of each channel Note that for such far-field photodetection based data collection schemes, the useful minimal dimensions of glass channel arrays will be determined by the resolution limit, i.e. related to the wavelengths of the light used.

In the second detection scheme used with channel array glass parallel separation means, molecular species are detected as they exit separation channels. Here, the output port comprising surface of said channel array glass may be briefly contacted with a receiving surface, such as an ion exchange paper surface or nylon membrane, onto which sample molecules are transported by the action of applied electrical field (or fluid flow in the case of ion-exchange border matrix capillary modifications described herein) and are bound to said receiving surface (e.g. by crosslinking with UV radiation in the case of nylon membranes.) Said contacting is performed such that sample molecule species bind to the portion of said sample receiving surface immediately juxtaposed to the channel from which each such molecule exited (i.e. such that sample molecules of an individual molecular species do not diffuse to regions juxtaposed to channels from which they did not respectively exit). This may be facilitated by performing said contacting with said receiving surface under high pressure normal to the juxtaposed surfaces, and may further be facilitated by the use of elastomeric surfaces or underlayers to cause said receiving surface to conform to said channel array glass surface comprising exit ports. Such brief contacting steps are repeated with fresh receiving surfaces as separation proceeds, preserving information as to the order in which receiving surfaces or receiving surface regions were contacted. Said receiving surfaces are then examined for the spatial location or pattern of labels, from which information about the sample sequences separated in each respective channel is reconstructed. Said pattern may favorably be captured by imaging visual labels (e.g. fluorescent labels, affinity labels bound by receptor-comprising beads, etc.,) with photodetector arrays such as CCD cameras, and analyzed by digital computer for information content, which is then reordered and further analyzed to glean sequence information.

It should be noted that various treatments of such channel glass arrays may further facilitate their use as means for parallel electrophoretic separations. First, though R.J. Tonucci et al.[38] do not note this, decreases in channel definition resulting from interdiffusion of the two or more glass types during drawing and annealing steps entail increases in internal surface roughness of the capillaries thus formed, when such interdiffusion is not so severe as

to completely interfere with channel formation. Surface roughness of internal channel surfaces will affect the separation properties of such channels in possibly favorable ways, which effect will depend partly on the internal volume to internal surface area ratio describing such channels.

A second category of treatments depend on the chemical modification or derivatization of said internal surfaces. As reviewed by A. Ulman,[39] the self-assembling monolayer systems based on organosilane compounds and derivatives thereof (e.g. $RSiX_3$, R_2SiX_2 or R_3SiX where R groups may be identical or distinct), which form covalent linkages to surface hydroxyl groups (e.g. formed on surfaces with exposed SiO_2 including native silicon oxides) may provide control over surface properties according to the particular compounds chosen. Other chemistries may similarly be used to derivatize such surfaces. This fact has been exploited in capillary zonal electrophoresis[40]. Derivatives conventionally include polyacrylamide derivatives, though these have generally been used for the reduction of protein adsorption to such surfaces. The embodiments of the present invention availing capillary tubes and channel array glass may further avail similar treatments. In particular, such treatments may provide for a channel comprising a surface localized separation matrix (border matrix) which may occupy a substantial portion of the inner diameter of such channels. Such a border matrix may comprise a predetermined polymer chain length or length distribution or composition distribution.

In particular, for the channels of channel array glass, or of capillary tubes, or of microfabricated separation channels, ionizable chemical functional groups situated on such border matrices may be used to modulate the degree of interaction of polynucleotides with such matrices, in analogy to ion-exchange resins. For example, pH of a running buffer (which occupies such channels and may flow through these independently of any border matrix) may be chosen to effect the desired degree of border matrix ionization, up to the maximum determined by the surface density of ionizable groups. Other conditions, such as temperature, affecting the ionization of chemical groups may similarly be exploited. Polynucleotide fragments comprise negatively charged phosphate linkages in direct proportion to length; for separation of polynucleotides, said ionizable groups are chosen

to yield cations upon ionization, in order to bind said negatively charged phosphate linkages. Degree of border matrix ionization will affect the probability that each negatively charged phosphate group associates with an ionized group of said border matrix. The different number of anionic groups corresponding to the different polynucleotide molecules will result in a difference degree of association of different length polynucleotides to such surfaces. Conditions and border matrix composition (and hence ionizability and ion exchange properties) are favorably chosen such that each polynucleotide fragment in the size range to be separated has an appreciable characteristic mobility, fractional association time or holdup, arising from differences in the mean number of associations with said border matrix per unit length. Optimal conditions and border matrix compositions will vary according to border matrix composition and size fractionation range, and the many parameters which may be varied to effect a desired result provide numerous possible combinations for many desirable fractionation properties. The mechanisms involved in these types of size fractionation may, for many cases, be envisaged as the association at a few points of polynucleotide backbones with a border matrix, where a characteristic proportion of said associations spontaneously dissociate per unit time, and said polynucleotide molecules thus *locally* freed (at one or more previously bound points along the length of said polynucleotide molecules) translate, to the extent possible under the constraint of other points of association persisting at that moment, in the direction of an applied electrical field or fluid flow. Thus, over time, a polynucleotide (or more generally, any polyions or polyanions subjected to such separation) advances in the direction of said applied field or fluid flow, at a rate which is negatively correlated with length (or more specifically, for polyions or polyanions in general, with number of cationic and/or anionic groups.) In the limit of less than on average one association per unit time, mobility will depend on the ratio of time which any such ion comprising species subjected to such separation spends in association with a border matrix ion to the time which it spends unassociated, and further depend on the mobility of said ion comprising species when unbound and in either the border matrix and overlying fluid phase (and the fraction of time spent in each of these, and mobilities in

each of these.) When transport in such a separation is directed by an electrical field, such methods may be termed ion-exchange resolved electrophoresis (or ion-exchange zonal electrophoresis), and when transport is directed by fluid flow such methods may be termed fluid injection directed ion exchange fractionation. Such methods bear some similarity to thin-layer chromatography, but are crucially distinguished by the presence of a free-fluid or mobile phase constrained to a limited proximity to said border matrix by channel dimensions.

Note that for any of these surface modification methods, particularly for surface modification aimed at providing surfaces suitable for ion-exchange based resolution methods, binary masking methods used with physical masks may be used to pattern surface density gradients or other distribution patterns of, for example, molecules comprising ionic or ionizable groups for separation methods involving ion exchange, where a surface used for forming a formed channel is exposed. Such a surface is a flat surface comprising linear trenches. Another such surface is a substantially flat surface which may be juxtaposed to said flat surface comprising linear trenches. Further, a large number of surface modification compositions and combinations thereof may be tested by such binary masking means for any desired sample type. This provides a method for the rapid optimization of surface modifying resolution media.

Combinations Including Partial Demultiplexation, Multiplex Separation or Parallel Detection:

<u>Partial Demultiplexation and Parallel Multiplex Separation:</u>

This aspect of the present invention combines partial demultiplexation, parallel multiplex separation and conventional serial reprobing of sequences separated in multiplex and transferred to membranes or the like. Parallel formed separatory channels capable of being opened length-wise may be used to enable transfer of separated species to membranes for probing. Alternatively, transfer to a moving membrane from the outlet port of capillaries or separatory channels simultane us with separation may accomplish transfer of the fractionated sample conserving a separation

coordinate (i.e. the time of exit coordinate is thus transformed to a positional membrane transfer coordinate on said membrane, related by the velocity of said membrane relative to said capillaries or separatory channel). The principal advantages of this aspect of the present invention are a considerably reduced requirement for probe arrays (requiring only the demultiplexation array), and the combination of both highly parallel separations and the rapidity of separation in small channels, with separation in multiplex. In this instance, each serial reprobing cycle yields data up to a maximum equal to read length times the number of separation channels used. For example, for a read length of 500 bases with 1000 separatory channels and a serial reprobing time (probing, washing, probe detection and probe removal) of one hour, 12 million bases of sequence information may be collected per membrane per day. Of course, reprobing of multiple membranes (i.e. patterns transferred from multiple sequencing runs) will increase this rate correspondingly. Using the results of E.S. Yeung et al. of 18 minutes per fractionation, 80 membranes may be generated in one day. This will be practical for a separatory channel pitch of 100 microns and in the case of longtitudinal transfer, length less than 50 cm. It should be noted that such a method will require reduced use of a parallel separatory channel array corresponding to the number of serial reprobings performed. Automated membrane transfer and serial reprobing equipment such as has been described may be adapted to embody this aspect of the present invention.[41,42,43] Because of the small quantity of each sequence species in each fraction or band, a preferred probe oligo labeling method in this aspect of the present invention is involves labeling, (e.g. with biotin) and detection with dye labeled or fluorescent labeled affinity modified beads (e.g. streptavidin modified fluorescent latex beads). In this case, intensity encoding with four monochromatic levels in one lane (referring to a region of the membrane) or channel is the preferred means for reducing the number of channels required. Data capture is again performed with a CCD and digital computer.

<u>Partial Demultiplexation, Parallel Multiplex Separation and Parallel Detection:</u>

Because both demultiplexation and parallel detection of species separated in multiplex both rely on the binding of tag sequenc s, it will be explicitly noted that for these variations of the present invention, samples must be constructed or provided such that the identity of a clone or sequence is specified by at least two distinct tag sequences, though said at least two distinct tag sequences may be contiguous or even overlapping, and may be naturally occurring, generated randomly or designed deliberately. The first of said tag sequences is a demultiplexation tag, and the second of said tag sequences is a detection tag for probing in a parallel detection step.

This aspect of the present invention gains maximum multiplexation of sample preparation and of sequencing reaction multiplexation, exploits parallel separation and further exploits parallel detection of multiplexed species in each separatory channel, simultaneously (i.e. a further level of parallelism) for different separatory channels. The maximum number of sequence reads which may thus be obtained is equal to the degree of parallelism of partial demultiplexation (which equals the number of separatory channels used) times the number of parallel detection elements for each fraction. Variations on and methods and means for this aspect of the present invention are described below.

Parallel Multiplex Separation and Parallel Detection:

These aspects of the multiplex loading and parallel separation methods and means of the present invention are combined with the parallel multiplex methods described by E.M. Rabani, comprising the simultaneous detection of plural different species within the same separation fraction. Such methods and means permit the sequence determination rate possible with a given separation rate (r) and a given separatory channel parallelism (s) to be multiplied by the parallelism of detection (d), yielding a maximal rate of (r)(s)(d) determined bases per unit time. This is favorably accomplished by lateral transfer, which is to say transfer of all fractions in one step or substantially simultaneously to detection arrays in a manner such that each fraction or each portion of a fraction will be

transferred to a distinct detection array or distinct detection array region.

Lateral Transfer of Fractionated Multiplex Species to a Second Surface or Detection Array:

In many instances it may be desirably to avoid the need for continuously moving parts and other complexities associated with the transfer of eluted fraction from exit ports of laterally sealed separatory channels to a moving membrane. The drawbacks associated with such methods may be avoided by using separatory channels which may be laterally opened or have many lateral ports, such that transfer may occur laterally as in the case of transfer from slab gels to laterally juxtaposed or opposingly placed membranes in, for example, Southern blots.

In the case of the present invention, lateral transfer from one or more separatory channel array facilitates the use of parallel detection arrays, preferably comprising stripes of plural distinct tag-probe oligonucleotides or other probing compositions, with an array repetition of a similar pitch to that of said separatory channel array, such that a said stripe of each of said distinct tag-probe oligonucleotides or other probing compositions occurs for each separatory channel of said separatory channel array. Said detection array comprising stripes are aligned parallel to said separatory channels during transfer, such that the contents of any fraction in each of said separatory channels will be exposed, upon transfer to at least one, and preferably exactly one, repetition of said stripes of said detection array. Alternatively, said array may comprise stripes of fine pitch which will be oriented perpendicularly to said separatory channels, provided that the pitch with which a full repetition of said array is less than half the length occupied by a species of a particular length in said separatory channel, such that the molecules in each fraction (many of which may be identical to each other but different from most molecules in the multiplex fraction) will be juxtaposed to all detection array elements.

In this aspect of the present invention, separatory channels are preferably formed from molded r cast microreliefs or reliefs (first reliefs), preferably of elastomeric composition, comprising narrow parallel trenches, which define the ultimate separatory channels, and

with surfaces composed or treated so as to have good affinity to any resolution media used but no undesirable affinity to sample molecules (in cases of separation by a gel or low viscosity medium, as opposed to separation on border matrices, in which case the surfaces of said reliefs are treated accordingly.) Said reliefs are juxtaposed to either a substantially flat surface or a second relief surface comprising arrays (fraction transfer arrays) of fraction transfer grooves and electrodes for lateral electroelution, and preferably also integrated tag-probe arrays; such juxtaposition is followed by the application of sufficient normal pressure to compress regions of greatest extremity on said relief or reliefs (or in the case of non-elastomeric relief surface, of an elastomeric film or an elastomeric second relief surface comprising arrays of fraction transfer groves, which is used as the opposed surface) so as to cause complete lateral sealing of said trenches into separatory channels by compression of the elastomeric material.

Fluid separatory media such as low viscosity media including polyethylene-oxide or polyacrylamide may then be injected into the resulting channels. Alternatively, in cases where said channels are formed by the juxtaposition of said first relief comprising trenches with a flat surface, a liquid resolution media (from which bubbles have been removed, e.g. by degassing) may be spread uniformly on one of the two surfaces, preferably said flat surface, which is in this case preferably flexible, which is applied with a uniform unrolling motion or a squeegeeing motion such that only one meniscus forms and is evenly translated so as to prevent the formation of bubbles. By such means, the need to inject resolution media into small channels may be avoided.

Transfer may be favorably accomplished as follows. Subsequent to separation of plural different multiplex samples on said formed separatory channel array, a surface of said formed separatory channel array comprising lengthwise portions of the internal surface area of each of said separatory channels is removed, and later replaced with a transfer surface, which is thus juxtaposed to the surface comprising the remainder of said formed separatory channel array. Thus, lateral opening of said channels is performed by removing, for example, said flat surface from said relief surface, the aspect ratio or recession of the trenches therein favoring retention of any

resolution media. This removal step is performed, preferably with a lateral sliding motion to prevent adhesion of resolution media, such that resolution media and sample material substantially remains in the remainder of said formed separatory channel array, principally or exclusively within the trenches defining formed channels.

After removal of said surface of said formed separatory channel array comprising lengthwise portions of the internal surface area of each of said separatory channels, said relief surface comprising trenches Said transfer surface is substantially free of separation media before juxtaposition to said surface comprising the remainder of said formed separatory channel array.

Alternatively, in cases where said formed separatory channel array comprised said second relief surface comprising arrays of fraction transfer grooves and electrodes for lateral electroelution, and preferably also integrated tag-probe arrays, separatory channels may be performed by juxtaposition, as in the other case, but further with relative positioning such that said trenches of said first relief are immediately opposite a flat portion of said second relief surface. Said second relief is preferably coated with a degassed liquid layer to facilitate electroelution before said juxtaposition is performed, and the resulting meniscus is caused to translate evenly to avoid the formation of bubbles. After fractionation, said second relief surface is translated by a lateral sliding motion along the surface of said first relief such that two openings of each of said fraction transfer grooves (which preferably run perpendicularly to said formed separatory channel trenches) are opposite each trench on said first relief which defines said separatory channel. After said lateral sliding causes the proper alignment of separation and fraction separation channels, a voltage is applied to the electrodes on of said second relief so as to electroelute fractions from said separatory channels and cause molecules in said fractions to pass into a restricted volume contacting said detection array. After said electroelution, said second relief may be translated so as to minimize the opportunity for denaturants to seep from said separation medium in said separatory channels into said restricted volume. Conditions are then changed to establish the desired binding stringency and thus permit selective binding of tag sequences with the tag-probes of said detection array.

After lateral transfer of multiplex fraction to said parallel detection arrays (which occurs in a stratified mann r) appropriate means are used for signal amplification, if any, and detection. For example, such detection methods and means may include contacting said parallel detection array to receptor or antibody labeled colored or fluorescent beads targeted to particular affinity labels used to label sequencing reaction products, e.g. according to terminator moiety. The positions in said parallel detection array and the associated fractionation coordinate are detected with suitable photodetection means, for example comprising a CCD array or photon camera and appropriate filters, and recorded by a digital computer interfaced to said photodetection means. Information concerning the sequence of sample material in communication with a particular second tag is reconstructed from the occurrence and identity of labels bound to the corresponding tag-probe of said tag-probe array in the ordered set of fractions, in a manner otherwise similar to sequence information reconstruction with other capillary gel electrophoresis sequencing methods. Said photodetection means and said detection array are preferably mounted on a system or translation table which moves one or both of these relative to each other such that data may conveniently be collected across the entire array by means of such relative motion or scanning. After data collection, said detection array is subjected to denaturing conditions and repetitive rinses to remove labels and sample molecules and thus prepare said detection array for reuse, care as always being taken not to contaminate said detection array.

Note that detection of samples on surfaces is facilitated by the signal amplification methods thus possible, including the aforementioned affinity label detection and signal amplification comprising the use of receptor labeled fluorescent beads, which methods are not favorably used with laterally sealed capillaries or gels.

Partial Demultiplexation, Parallel Multiplex Separation and Parallel Detection:

The same methods and means permitting the application of multiplex samples above to an array of separatory channels may be combined with the above described embodiment of the present invention, such that

partial demultiplexation by a demultiplexation array sequences sample material, preferably subsequent to sequencing reactions, targeting molecules comprising particular first tag sequences to corresponding particular individual separatory channels of said array of separatory channels, such that identity of said first tag (more specifically the tag or tag sequence utilized for demultiplexing, as distinguished from a second tag sequence used for parallel detection) determines the channel to which each molecule will be transferred. By these methods and the corresponding combined means, a large multiplex sample or sets of multiplexed samples may be applied to parallel separatory channels or separatory channel arrays in a single step, in preparation for the parallel multiplex separation and parallel detection for each channel described above.

Determination of Tag Sequences Not Present in a Complex Sample:

For many large scale sequencing applications of the present invention, it may be desirable to determine the sequences of a large number of candidate tag (and probe-tag) sequences suitable for use with that sample, where each such tag sequence occurs infrequently or not at all within a sample. One method for accomplishing this is as follows. A pool of oligonucleotides of a particular nucleotide length n is randomly synthesized from all for nucleotide monomer reactants by a suitable oligonucleotide synthesis chemistry (e.g. phosphoramidite). Length n must be chosen such that number of different sequences in said pool, 4^n, exceeds the size N (number of bases) of the sample by some factor greater than, for example, two. A sample of size N can comprise no more than N distinct n-mer sequences (N>>n). Thus, said pool comprises a large number of n-mers not present in said sample. The non-randomness and repetitivity of evolved DNA sequences increases this number.

We may consider the example of a genome of $N=10^9$ bases in size. Such a genome is of lower complexity than a pool comprising all possible 15-mer oligonucleotides. We may thus choose a pool comprising half of all possible 16-mers, by reducing the number of nucleotides (i.e. to <4, e.g. to two) at one position of said n-mers. Such a pool is produced in quantity sufficient to ensure that on average at least, for example, ten thousand copies of each possible

sequence occur. At least half of this quantity is left on the solid phase support used in synthesis but is nonetheless subjected to removal of protecting groups. A quantity equivalent to 100 copies of each sequence in the pool is fractionated against genomic DNA. This is equivalent to about 2ng or 0.4 pmol. This quantity of such an oligonucleotide pool is mixed with a tenfold excess of genomic DNA (which may have previously been filtered to eliminate small fragments from the sample). The resulting mixture is subjected to denaturing conditions and then hybridized under reduced stringency hybridization conditions. Reduced stringency conditions are favorable because of the variation optimal hybridization conditions for highest stringency with (oligo) sequence; reduced stringency conditions will eliminate sequences which, although they do not occur in said genome, nonetheless may bind to genomic sequences under the set of conditions used, while minimizing the likelihood that any oligos with sequences occurring in said genome will fail to bind. Unbound oligonucleotides are recovered after hybridization, for example by ultrafiltration (care being taken to avoid excessive fragmentation by shearing, etc.,). The unbound fraction may be subjected to repeated rounds of depletion by repetition of the foregoing mixing, denaturing, hybridization and filtration steps, against the same quantity of genomic DNA used initially, to greatly enrich for sequences not occurring in said genomic DNA and greatly reduce the presence of genomic sequences in said unbound oligonucleotide fractions. Different conditions may be used in different depletion rounds.

Following depletion against genomic DNA, the fraction unbound after the final depletion cycle may optionally then be subjected to selection for binding of the undepleted pool, under more stringent conditions. This is again performed by mixing the undepleted, solid phase bound pool in, for example, fifty-fold excess, denaturing and hybridizing under more stringent conditions. Washing or filtering is then performed to eliminate the unbound fraction. The bound fraction may then be recovered by denaturation and washing, and will appear in the eluted fraction. Such a selection will enrich for sequences which bind their complementary sequences under more stringent conditions. Such enrichment will reduce the number of oligo sequences not bound to genomic DNA but exhibiting unacceptable non-specificity of binding against the lager complexity random pool, due

to the increased stringency used. This treatment does not, however, formally preclude the possibility of recovery of oligos which bound to less than fully complementary targets because no discrimination of such events is possible, and these are only reduced with increasing stringency. Further, binding of oligos in this fraction to more than one tag-probe cannot be excluded, but will be improbable under sufficiently stringent conditions. Longer random tag sequences (relative to a fixed sample genome complexity) prepared as above and used under correspondingly higher stringency binding conditions in sequencing protocols and other related protocols may address this issue by rendering the likelihood of one probe-tag binding multiple tags or one tag binding to other tags correspondingly less likely in finite sample sizes. Further, to the extent that such events occur, these will only reduce overall efficiency of parallel detection or demultiplexing across a large sample diversity (e.g. samples comprising many different distinctly tagged sequence inserts); these inefficiencies are readily offset by the parallelism and multiplexing possible with the methods of the present invention.

The fraction thus recovered comprises sequences favorably used as tag-probes, sequences complementary to these may favorably be used as tags for library clone tagging. Sequence determination of the oligos obtained by the above methods may be performed as follows. Two 5′ phosphorylated oligos, complementary to each other over most of their length but having, for example, a random three-base 3′ overhang at one end (but blunt at the other) are annealed to an excess of depleted tag-probes, and then ligated to these, such that a ligation event occurs at the 3′ end of oligos from the pool of said depleted tag-probes. The resulting products are then subjected to polymerization (primer extension) to fill in the overhangs thus formed comprising 13 of the 16 bases of said depleted tag-probes. Finally, the resulting mixture is subjected to blunt-end ligation (e.g. with T-4 DNA Ligase) to form large multimers, which are then ligated into a sequencing vector, or a vector comprising priming sites suitable for PCR amplification of the insert material. These vectors are then transformed into hosts, or the inserted material is amplified by PCR after dilution and distribution into wells (e.g. of a microtitre dish suitable for PCR cycling) at a concentration yielding on average one amplifiable molecule per well. Vectors used

favorably comprise a second tag sequence immediately downstream from a standard sequencing primer sequence and immediately upstream from the inserted material. Said second tag sequence comprising a limited number of variable bases at predetermined positions, e.g. 6 bases, not necessarily all contiguous, such that probe arrays complementary to all of the possible second tags may be formed and used for one or more of the demultiplexing or parallel detection methods of the present invention. A third tag sequence of predetermined complexity may similarly be included, for example, adjacent to said second tag sequence, and may comprise a similar complexity but a distinct consensus sequence, and be used in a step other than where said second tags were used for demultiplexing or parallel detection, said third tags used for parallel detection or demultiplexing. Thus, a set of tags from the depleted pool may be sequenced, and the methods of the present invention may be used to rapidly determine the sequences thus recovered, where said second and said third tags permit, for example, demultiplexing (here, partial demultiplexation) for separation of a complex pool in parallel separatory channels by binding to arrays comprising probes for said second tags, and parallel detection is accomplished with arrays of probes complementary to said third tags. Sequence information gained about sequences not occurring with such a genome is favorably then used to generate probe tag arrays for used with said genome. Further, the cloned or amplified depleted tag material may be constructed such that individual tags are liberated by digestion with a predetermined restriction enzyme, and these may then be used in the generation of a pool of tagged derivatives of a library vector construct. Alternatively, sequence information regarding said depleted tag material may be used to score consensus sequences and m-mers ($m<n$) which may be used as elements in the synthesis of pools of oligos comprising sequences of highly unlikely occurrence in said genome.

For tagging downstream from a primer, in which case a standardized primer of identical sequence may be used with all clones from such a library such tag sequences are preferably placed within a library vector immediately downstream from a sequencing primer region, and immediately upstream from a cloning site. Thus, such tag sequences will be copied (or otherwise included) during the generation of

sequencing reaction products, and all but the shortest sequencing products will comprise the tag sequence.

Constructs related to the vectors described above are disclosed by G.M. Church and S. Kieffer-Higgins in U.S. Patent Number 5,149,625. Such constructs may similarly be used with certain various aspects of the present invention. Where tag-probe arrays are to be used for demultiplexing or parallel detection, said probe-tag array oligonucleotide ensemble composition must comprise oligonucleotides having sequence composition which will generally, and preferably uniformly, each probe a distinct tag of such vectors under sufficiently stringent binding conditions. Thus, one aspect of the present invention is the demultiplexation based separation of libraries or other pools of different sequences constructed from the vectors of G.M. Church and S. Kieffer-Higgins.

Copolymer Array and Probe-Tag Array Patterned Synthesis:

Synthesis Patterned by Lithographic-Plate:

Lithographic plates such as those taught by T. Suzuki and F. Shinozaki in U.S. Patent Number 5,380,620, or those reviewed by B. Bednar, J. Kralicek and J. Zachoval,[44] which accomplish the spatially predetermined transfer of liquid patterns, where said liquid is conventionally an ink, colorant or etch-resist, may be applied to the fabrication of copolymer arrays. Copolymer types may include peptides and oligo- and poly-nucleotides, as well as other copolymers which may be synthesized by deprotection step-controlled chemistry on solid phase supports. This method is in analogy to that of S.P.A. Fodor et al., but substitutes spatial control of reagent placement according to the wettability of an juxtaposed surface contacted with an array substrate for spatial control of removal of photolabile protecting groups. The principle advantages of this present method is that monomers (comonomers, e.g. phosphoramidite deoxynucleotides of each of the four nucleobase compositions found in naturally occurring DNA) need not be chosen according to photodeprotection requirements, which are eliminated, and the use of light or actinic radiation exposure equipment is also eliminated. Thus, standard p ptide synthesis reagents or polynucleotide (e.g. DNA) synthesis reagents and comonomers may be used. Three variations are possible.

In a first variation, lithographic plates are used to contact regions of said array, corresponding to those regions desired to undergo deprotection of the copolymers or initiators thereupon, with deprotection reagents included in said liquids ("inks") on said plates. The entire surface of said array is then exposed to the reactant (comonomer) to be added only to the copolymers or initiators deprotected by the preceding treatment. By such means, the addition of comonomers to only those copolymers or initiators situated in a spatially predetermined set of regions (according to the patter and relative positioning of the lithographic plate used in the immediately preceding step) is effected. Such steps are repeated, using different lithographic plates aligned relative to the position of precedingly used lithographic plates, in an overall procedure otherwise resembling the light controlled method of S.P.A. Fodor et al., to produce the desired pattern of copolymers with predetermined sequence composition and situated in a predetermined manner which corresponds to said sequence composition on said array substrate surface. This first variation may be termed lithographic plate patterned deprotection. A second, contrasting, variation may be termed lithographic plate patterned comonomer contacting and addition. Here, the totality of one or more such array is contacted with deprotection reagents to remove any protecting groups from any preceding steps, and then contacted with a lithographic plate "inked" with a liquid comprising a single comonomer reactant type. Comonomers are thus added only to those copolymers or initiators in those array regions contacted with said liquid, i.e. contacted with said reactant. With this method, a wash step is favorably performed using reagents that will inactivate unadded comonomer reactants, such that the reactivity of these is neutralized before they are swept into possible contact with regions comprising copolymers or initiators with which said reactants are not desired to react. The advantage of this variation lies in the reduction of steps from the identity of patterning and reactant contacting steps. As in any patterned copolymer synthesis method, the final pattern is determined by the lithographic plate pattern used, the order in which these are used, and the alignment and precision of alignment with which they are repeatably contacted with said array substrate. The third variation combines the patterning methods of th previous variations,

and may be termed lithographically patterned deprotection with redundantly patterned comonomer contacting and addition. Here, lithographic-plate control is exerted, as in the first variation above, to effect the deprotection of initiators or copolymers situated in spatially predetermined array regions in a first cycle step, and in a second cycle step either the same lithographic plate or another plate with an identical pattern is used, as in the second variation above, to contact reactants with only the desired predetermined regions substantially identical in location and extent to those deprotected in said first cycle step, after which a wash step is performed as in the second variation above. Such a cycle is repeated with different lithographic plates to produce the desired final copolymer array pattern.

The large areas which are readily patterned by such lithographic means make these methods particularly advantageous in the high volume production of copolymer arrays, particularly arrays of extended dimensions. It must be emphasized that the patterning method disclosed here relies primarily upon the differential wetting of surface regions of said lithographic plates rather than any other principle of liquid pattern formation and transfer.

These methods may be availed to produce oligonucleotide arrays useful for sequencing by hybridization, for polynucleotide characterization, for the determination of overlap of clones in a DNA library, for mutation or genetic disease detection, or for to produce the probe-tag arrays used in the parallel multiplex sequencing method previously described by E.M. Rabani cited above. These methods, as well as any other functionally adequate or comparable methods, including the photodeprotection methods of S.P.A. Fodor et al., or those taught by J.L Winkler, S.P.A. Fodor et al., in U.S. Patent Number 5,384,261 comprising steps of flowing reagents through groves on the surface of a channel block juxtaposed to a substrate, or those taught by R.E. Southern and co-workers in U.S. Patent Number 5,324,663, may be employed to produce the linear probe-tag arrays disclosed herein as useful for parallel multiplex sequencing with direct spatial transfer including parallel multiplex sequencing with parallel separation and direct spatial transfer.

Note that the lithographic plates used in this aspect of the present invention may have patterns comprising isolated inked or

isolated wettable regions, as retention of applied solutions (rather than transport through channels to desired surface areas, as in some prior art methods) does not require transport through contiguous regions. Note further that lithographic plates of this kind are favorably produced by the microfabrication methods described and referenced above, and thus generally comprise films of polymeric resist-type coatings which are usually of 50 nm or greater depth from the surface, being deposited by such methods as spin-coating.

In summary, this method comprises the steps of: (a) producing lithographic plates that specify the desired spatial patterns of chemical reactions, in particular, the desired spatial patterns of addition of reagents or reactants; (b) applying said reactants or reagents to one of said lithographic plates for use in a reaction cycle according to the final pattern desired, with the result that solutions are only retained in selected predetermined regions according to the pattern of the respective lithographic plate; (c) juxtaposing the lithographic plate of step (b) to the target surface on which copolymer initiators or intermediate products are situated such that reagents or reactants are contacted and thus permitted to react with only the desired regions specified by the pattern of said lithographic plate, where said juxtaposing constitutes opposing said lithographic plate to said target surface in correct alignment, register and relative position (e.g. relative to some reference point which facilitates the correct registry of material applied by different lithographic plates); (d) performing any further necessary chemical reactions on said target surface which do not require spatial limitation; and (e) repetition of steps (b) through (d) as many times, and with lithographic plates of correct pattern, as necessary for the desired copolymer array pattern to result. As will be obvious to those skilled in the relevant arts, such a method will be suitable for any copolymer which may be produced by a synthesis procedure wherein some controlling step is performed with a solvent or reagent for which there can be found a material composition to yield a surface which repels said solvent or reagent and for which there can be found a materiel composition to yield a surface which retains said solvent or reagent. Copolymers of particular interest which may be produced by this method are DNA, RNA, peptide-nucleic

acids, peptides and the various peptide analogues or mimics, and obvious variations upon all of these.

Copolymer and Modified Oligonucleotide Synthesis by Patterned Activation with Metal Catalyst Structures:

Probe-tag arrays, arrays comprising oligonucleotides which need not be probe-tags, and most generally copolymer arrays may be produced by patterned synthesis comprising one or more steps of spatially controlled catalysis, particularly spatially controlled deprotection or activation. One aspect of this method, comprising the catalytic reduction of nitrile groups in the presence of dissolved H_2 to yield amine groups will be presented here for illustration. It will be noted, however, that the essence of this method concerns the exertion of spatially predetermined step control in the synthesis of copolymers situated on surfaces, and may be adapted to other chemistries and corresponding reactions catalyzed by metal surfaces.

For the synthesis of backbone modified oligonucleotides, nucleotide synthesis reactant monomers comprising 5′ nitrile groups, and not comprising phosphoester protecting groups which include nitriles (such as is the case for cyanoethyl groups used with phosphoramidite oligonucleotide synthesis chemistries) or any unsubstituted amides are used. The active acceptor group in preferred reactant monomers is a 3′ -($OPOCH_3Cl$) activated phosphate. This group is chosen to be reactive with amines (by attack of amine nitrogen lone pair at the partially positive phosphate, with liberation of the chlorine as a stable anionic leaving group), and to not comprise any nitrile or unsubstituted amide functionalities. Initiators used, which are preferably of substantially linear structure, comprise a surface binding group (e.g. an organosilane for glass substrate surfaces), a linker of sufficient but not excessive length (e.g. having linkers 10-1000 atoms or bonds in length), and a terminal nitrile group. Longer initiator linkers will impose limitations on minimum array element size and spacing. Like azide groups, nitrile groups may be reduced by platinum in the presence of dissolved H_2.[45]

Patterned platinum catalyst structures or microstructures, which may, for example, be patterned and fabricated by LIGA methods,

comprising substantially flat surfaces from which a pattern is recessed or omitted (in the form of voids), provides for the spatially predetermined patterned activation of nitrile groups situated on a surface. The patterned surface of said patterned platinum catalyst structure is contacted, in the presence of dissolved H_2, with said substrate surface on which said nitriles are situated; said nitriles experiencing sufficient contact with the surface of said platinum are thus transformed to amine groups. Said contacting is performed in a manner and with means such that a particular platinum catalyst structure may be repeatably contacted with the same surface with high spatial accuracy, or with high accuracy stepping (as in semiconductor microlithography). The entirety of said substrate surface is then contacted with nucleotide reactant monomers such as those described above comprising the base moiety (in protected form) desired to be added to the amines formed in the preceding step; nitriles are substantially inert to methoxyphosphorochloride ester addition, as evidenced by their use as backbone protecting groups in other oligonucleotide synthesis chemistries employing PCl_3 derivatives. A catalytic patterning step followed by a monomer addition step comprise a patterned addition cycle. Other steps necessary are identical to those used in standard phosphoramidite chemistries; these include capping and oxidation steps. Patterned addition cycles are repeated with different patterned platinum catalyst structures or microstructures, or the same structures in different registry or juxtaposed orientation, and different reactant monomers such as those described above comprising different base moieties each of which said moieties is desired to be added to the amines formed in the preceding catalytic patterned activation step. The oligonucleotide synthesized by nitrile/amine step control methods such as these will comprise backbone linkages replacing 5'CH_2-(NH)-PO_2-O-3'C for the usual 5'CH_2-O-PO_2-O-3'C. Under conditions where this backbone nitrogen is not further protonated (i.e. under mildly basic conditions), hybridization properties of oligonucleotides with such modifications are expected to be similar to those of unmodified oligonucleotides of identical sequence. The logic and considerations pertaining to the spatial patterns chosen and order of comonomer addition to yield an array with a desired oligonucleotide array sequence compositional

configuration are substantially identical to those pertaining to the methods of S.P.A. Fodor et al.,[46] or other similar methods.

Note further that such methods could activate regions of surfaces to the immobilization of small-molecule haptens or ligands, for use in production of arrays useful for parallel assays for the determination of affinity or binding of receptors, immunoglobulins, enzymes, cells, ribozymes or other macromolecules. Likewise, biomimetic copolymers which may be prepared from monomer precursors comprising nitrile groups may likewise be thus prepared.

Copolymers of particular interest which may be produced by this method are backbone-modified-DNA, backbone-modified-RNA, peptide-nucleic acids, peptides and the various peptide analogues or mimics, and obvious variations upon all of these.

<u>Copolymer and Modified Oligonucleotide Synthesis by Patterned Activation with Metal Catalyst AFM Tip Surface:</u>

The same synthesis methods used above with patterned metal catalyst surface controlled activation patterning may be used to synthesize copolymer arrays with nanometer resolution by the use of scanning probe positioning means (e.g. a scanning probe microscope) to position a catalyst metal coated probe (e.g. an AFM tip), as was done by W.T. Müller, P.G. Schultz, et al.[47] but otherwise according to the above oligonucleotide synthetic methods. In other words, activation of initiator nitriles or nucleotide or oligonucleotide nitriles, to form amines which will react as donors with phosphate derivative groups on nucleotide monomer reactants is performed here by the positioning and contacting of the platinum on an AFM or other similar tip with said substrate surface, but synthesis is otherwise conducted as above.

Arrays used with the present invention may be used to bind polynucleotide species labeled with one among several different affinity labels, which affinity labels may be detected with beads comprising receptors or antibodies or other binding means specific for each of said different affinity labels. Each of said receptors or antibodies or other binding means is placed in communication with a perceptibly distinct bead (e.g. streptavidin with red beads and antidigoxigenin with yellow beads).

Note that such tag-probe arrays may be synthesized on high dielectric, low dielectric or conducting materials with appropriate initiators (e.g. on Au(111) with linkers comprising terminal -SH groups) or on membranes of polymeric composition (e.g. nylon, latex, etc.,) chosen or treated so as to be inert to synthesis conditions and reagents, where transfer of species from a gel by fluid absorption (as with the transfer of polynucleotides from gels to membranes in Southern and Northern blots) is desired, which choice will be determined by the desired method of transfer of fractionated species in each separation channel of a separation channel array to said tag-probe arrays. Probe-tag arrays situated on surfaces of such compositions favorable for blotting transfer may be produced by the methods and means disclosed herein, by those of known art including those photolithography based methods of S.P.A. Fodor et al. or any other equivalent means compatible with such surfaces, appropriate initiators and linkers being chosen.

The foregoing embodiments and examples have been presented for illustration rather than limitation, with the many possible variations, changes and alternative embodiments which will be obvious to those of ordinary skill in the relevant arts included within the scope and spirit of the present invention. Thus, the scope and the breadth of the present invention is intended to be limited only by the appended claims rather than the foregoing description;

---

[1]Maxam, A.M. ; Gilbert, W.; **1977**. *Proc. Natl. Acad. Sci. U.S.A.* *74*:560.

[2]Sanger, F.; Nicklen, S.; Coulson, A.R.; **1977**. *Proc. Natl. Acad. Sci. U.S.A.* *74*:5463.

[3]See, for example, a standard laboratory manual is: Sambrook, J.; Fritsch, E.F.; Maniatis, T.; **1989**. *Molecular Cloning: A Laboratory Manual*, Second Edition. Cold Spring Harbor Laboratory Press.

[4]For example, low viscosity media as taught U.S. Patent Number 5,374,527.

[5]Church, G.M.; Gilbert, W.; **1984**. *Proc. Natl. Acad. Sci. U.S.A.* *81*:1991.

[6]Cherry, J.L.; et al.; **1994**. *Genomics, 20*:68.

[7]Church, G.M.; Kieffer-Higgins, S.; **1992**. U.S. Patent Number 5,149,625.

[8]Church, G.M.; **1992**. U.S. Patent Number 5,149,625.

[9]Rickwood, D.; Hames, B.D., eds.; **1982**. *Gel Electrophoresis of Nucleic Acids: A Practical Approach*. I.R.L. Press, Washington, D.C.,.

[10]Sambrook, J.; Fritsch, E.F.; Maniatis, T.; **1989**. *Molecular Cloning: A Laboratory Manual*, Second Edition. Cold Spring Harbor Laboratory Press.

[11]Lagerkvist, A.; et al.; **1994**. *Proc. Natl. Acad. Sci., U.S.A., 91*:2245.

[12]This difficulty has been noted by both E.S. Yeung and L.S. Smith. See news article, Borman, S.; 1995. Chemical and Engineering News, 24 July 1995, p.37-39.

[13]Tonucci, R.J.; et al.; **1992**. *Science, 258*:783.

[14]Woolley, A.T.; Mathies, R.A.; **1994**. *Proc. Natl. Acad. Sci., U.S.A.,91(24)*:11348.

[15]Fung, N.E.; Yeung, E.S.; **1995**. *Analyt. Chem., 67(13)*:1913.

[16]Huang, X.C.; Quesada, M.A.; Mathies, R.A.; **1992**. *Analyt. Chem., 64(18)*:2149.

[17]Aumiller, G.D.; Chandross, E.A.; Tomlinson, W.J.; Weber, H.P.; **1974**. *Jour. Appl. Phys., 45(10)*:4557.

[18]Becker, E.W.; Ehrfeld, W.; et al.; **1986**. *Microelectronic Engineering, 4*:35.

[19]Abraham, M; Ehrfeld, W.; et al.; **1994**. *Proc. SPIE, 2213*:48.

[20]Gale, M.T.; et al.; **1994**. *Proc. SPIE, 2213*:2.

[21]Neyer, A.; et al.; **1993**. *Electronics Letters, 29(4)*:399.

[22]Klein, R.; Neyer A.; **1994**. *Electronics Letters, 30(20)*:1672.

[23]Bednar, B. ;Kralicek, J.; Zachoval J.; **1993**. *Resists in Microlithography and Printing*. Elsevier. See pages 289-294.

[24]A recent example of this work may be found in Wilbur, J.M.; Whitesides, G.M.; et al.; **1995**. *Advanced Materials, 7(7)*:649.

[25]Müller, W.T.; Schultz, P.G.; et al.; **1995**. *Science, 268*:272.

[26]McIntyre, B.J.; Salmeron, M.; and G.A. Somorjai; **1994**. *Science, 265*:1415.

[27]For example, see Beavis; R.C.; Chait; B.T.; **1994**. U.S. Patent Number 5,288,644.

[28]For example, see Jett, J.H.; Keller, R.A.; *et al.*; **1990**. U.S. Patent Number 4,962,037.

[29]For examples, see: Driscol, R.J.; Youngquist, M.G.; Baldeschwieler, J.D.; **1990**. *Nature, 346*:294;

Dunlap, D.D.; Bustamante, C.; **1989**. *Nature, 342*:204;

Allison, D.P.; *et al.*; **1992**. *Proc. Natl. Acad. Sci., U.S.A. 89:*:10129;

and for a review, see: Stine, W.B.; Rabani, E.M.; Smith, D.W.; *et al.*; **1995**. "Imaging of DNA Bases and Oligodeoxynucleotides using Scanning Tunneling Microscopy," in: *Frontiers in Biological Physics*. Academic Press, San Diego.

[30]Lindsay, S.M.; Philipp, M.; **1992**. U.S. Patent Number 5,106,729.

[31]Fodor, S.P.A.; et al.; **1991**. *Science, 251*:767.

[32]Bains, W.; **1994**. *Genome Analysis, Techniques and Applications, 11(3)*:49.

[33]Broude, N.E.; et al.; **1994**. *Proc. Natl. Acad. Sci. U.S.A., 91*:3072.

[34]Broude, N.E.; et al.; **1994**. *Proc. Natl. Acad. Sci. U.S.A., 91*:3072.

[35]Fodor, S.P.A.; et al.; **1991**. *Science, 251*:767.

[36]Lagerkvist, A.; et al.; **1994**. *Proc. Natl. Acad. Sci., U.S.A., 91*:2245.

[37]See, for example, Ford, J.E.; Esener, S.; et al.; **1993**? *Proc. SPIE, 2026*:604.

[38]Tonucci, R.J.; et al.; **1992**. *Science, 258*:783.

[39]Ulman, A.; **1991**. *An Introduction to Ultrathin Organic Films, From Langmuir-Blodgett to Self-Assembly*. Academic Press, Inc.,. See especially pages 237-304.

[40]Weinberger, Robert; **1993**. *Practical Capillary Electrophoresis*. Academic Press, Inc., p.56-59.

[41]Cherry, J.L.; et al.; **1994**. *Genomics, 20*:68.

[42]Church, G.M.; Kieffer-Higgins, S.; **1992**. U.S. Patent Number 5,149,625.

[43]Church, G.M.; **1992**. U.S. Patent Number 5,149,625.

[44]Bednar, B. ;Kralicek, J.; Zachoval J.; **1993**. *Resists in Microlithography and Printing*. Elsevier. See pages 289-294.

[45]See, for example, Roberts, J.D.; Caserio, M.C.; 1977. Basic Principles of Organic Chemistry, Second Edition. W.A. Benjamin, Inc.,. p. 1146.

[46]Fodor, S.P.A.; et al.; **1991**. *Science, 251*:767.

[47]Müller, W.T.; Schultz, P.G.; et al.; **1995**. *Science, 268*:272.

whereforth, I claim:

Claims:

1. Method for determining the sequences of plural polynucleotide samples comprising one or more of: manifold sample transfer; loading of two or more distinct separatory channels or means or regions thereof each comprising one or more surface immobilized oligonucleotide sequences at or near the input of said separatory channels or means or regions thereof and comprising a step of hybridizing said polynucleotide samples to said surface immobilized oligonucleotide sequences; loading of products resulting from the amplification of polynucleotide molecules randomly distributed into one or more bin arrays;
and further comprising one of: transfer of fractionated polynucleotide molecular species to a separate surface or a surface region not part of the surface of a separatory channel, detection probe array or membrane subsequent to separation or as separation is conducted; transfer of fractionated polynucleotide molecular species to a tag-probe array during or after fractionation; transfer of fractionated polynucleotide molecular species to a tag-probe array after fractionation by means of an array comprising fraction transfer grooves and two or more surface immobilized tag-probes and electrodes for lateral electroelution.

2. The method of claim 1 further comprising separation steps with separation means selected from the group consisting of: parallel capillary array, parallel separatory channel array, channel array glass, microfabricated separatory channels.

3. Method for determining the sequence of two or more different polynucleotide samples simultaneously according to the method of claim 1 comprising sample demultiplexing means situated at or near the entry port of each of two or more separatory channels.

4. Method for determining the sequence of two or more different polynucleotide samples simultaneously according to the method of claim 3 comprising two or more separatory channels which may be opened along their length and tag-probe or oligonucleotide arrays

comprising two or more different sequences with patterns aligned along or repeating periodically perpendicularly to the direction of said separatory channels.

5. Method according to claim 1 further comprising the use of optical detection means for detecting tag-probe array or parallel detection array bound species .

6. An automated device for sequencing polynucleotides according to the method of claim 5, comprising one of:

one or more manifolds for sample transfer; two or more distinct separatory channels or means or regions thereof each comprising one or more surface immobilized oligonucleotide sequences at or near the input of said separatory channels or means or regions thereof; one or more bin arrays for the amplification and/or transfer of single polynucleotide molecules.

and further comprising one of:

one or more separate surfaces or surface regions not part of the surfaces of one or more separatory channels, detection probe arrays or membranes, for transfer of fractionated polynucleotide molecular species subsequent to separation or as separation is conducted; one or more tag-probe arrays for the receiving of fractionated polynucleotide molecular species during or after fractionation; one or more tag-probe arrays for receiving fractionated polynucleotide molecular species after fractionation of said polynucleotide molecular species by means of an array comprising fraction transfer grooves and two or more surface immobilized tag-probes and electrodes for lateral electroelution.

7. Method for the parallel loading of parallel capillary electrophoresis apparatus comprising the use of one or more one- or two-dimensional comb or comb-like structures to individual teeth of which individual polynucleotide samples are reversibly bound or immobilized.

8. Method for the parallel loading of parallel capillary electrophoresis apparatus according to claim 7 where said one or more one- or two-dimensional comb or comb-like structures further comprise on said individual teeth, or on the surface thereof or otherwise in communication with said surface, an oligonucleotide probe such that said one or more one- or two-dimensional comb or comb-like structures comprises on its surface an oligonucleotide probe array, in configurational register with the array of said teeth, to which said teeth individual polynucleotide samples are reversibly bound or immobilized by hybridization with predetermined stringency.

9. Method for the parallel loading of a parallel capillary electrophoresis apparatus according to claim 7 where said one or more one- or two-dimensional comb or comb-like structures are employed to transfer sample species from one of: polynucleotide molecules resulting from the random distribution of polynucleotide molecules into a bin array, such that many bins commonly contain exactly one polynucleotide molecule, followed by a step amplifying said polynucleotide molecule in the bins of said bin array; the random distribution of polynucleotide molecules, such that many bins commonly contain exactly one polynucleotide molecule, into two or more bin arrays followed by a step amplifying said polynucleotide molecules in said bin arrays and a step selectively transferring, on an array element by element basis those products produced in said amplifying step arising from a single molecule in a single element of said two or more bins into the corresponding element of a third bin array or the corresponding channel of a separatory channel array, such that each said corresponding element of said third bin array or each said channel of said separatory channel array receives material amplified from exactly one molecule.

10. Method for the parallel loading of a parallel capillary array comprising the use of one or more oligonucleotide probe arrays the elements of which occur in a spatial configuration complementary to that of the input ports of said parallel capillary array, to which sample mixture of plural polynucleotide

molecule species are bound by hybridizati n in a first demultiplexation step, after which said one or more oligonucleotide probe arrays are then juxtaposed to the said input ports of said parallel capillary array, followed by a denaturation step to release said plural polynucleotide molecule species such that only said polynucleotide molecule species bound by a single said element of said oligonucleotide probe arrays are released into the single juxtaposed said input ports of said parallel capillary array, in a second transfer step, for subsequent electrophoretic separation and detection of the polynucleotide molecular species thus separated.

11. A method for the parallel synthesis of patterned copolymer arrays on a solid surface comprising the steps of:

(a) producing one or more lithographic plates, the surface of each comprising at least a first portion which is wetted by a reactant solution and a second portion which repels said reactant solution, and where said first portion or said second portion comprises and is defined by a polymeric coating which differs in composition from the underlying material of said lithographic plate;

(b) uniformly applying a first reactant solution to one said lithographic plate of step (a);

(c) juxtaposingly contacting said lithographic plate with applied reactant solution of step (b) to said solid surface, in proper register and alignment with said solid surface;

(d) performing any necessary chemical reactions necessary to the synthesis of the desired copolymers which need not be under spatial control to obtain the desired surface pattern on said solid surface;

(e) repeating steps (b) through (d) a sufficient number of times to obtain a surface decorated with a predetermined spatial pattern of two or more distinct copolymer sequences.

12. The method of claim **11** where said copolymer array comprises oligomers of composition selected from the group consisting of: DNA; RNA; peptide-nucleic acids; polypeptides; and, polypeptide mimics.

13. A method for the parallel synthesis of patterned copolymer arrays on a solid surface comprising the steps of:
(a) producing one or more catalytic surfaces the surface configuration of which is known;
(b) reacting initiator molecules to a solid surface;
(c) uniformly applying a first reactant solution to one said catalytic relief of step (a);
(d) juxtaposingly contacting one of said catalytic surfaces, in proper register and alignment with, and relative position to said solid surface, and under conditions favorable to catalysis and the desired chemical reaction, to said solid surface;
(e) performing any necessary chemical reactions necessary to the synthesis of the desired copolymers which need not be under spatial control to obtain the desired surface pattern on said solid surface;
(f) repeating steps (b) through (f) a sufficient number of times to obtain a surface decorated with a predetermined spatial pattern of two or more distinct copolymer sequences.

14. The method of claim 13 where said copolymer array comprises oligomers of composition selected from the group consisting of: backbone-modified DNA; backbone-modified RNA; peptide-nucleic acids; polypeptides; and, polypeptide mimics.

15. A method for the parallel synthesis of patterned copolymer arrays on a solid surface comprising according to claim 13 where said catalytic surfaces, which are of known configuration, are catalytic reliefs comprising predetermined surface patterns.

16. A method for the parallel synthesis of patterned copolymer arrays on a solid surface comprising according to claim 13 where said catalytic surfaces are the surfaces of one or more atomic force microscope or scanning probe microscope tips.

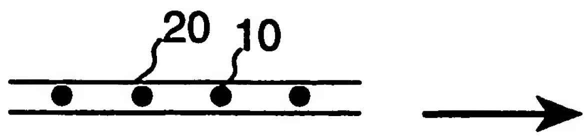

Fig. 1a

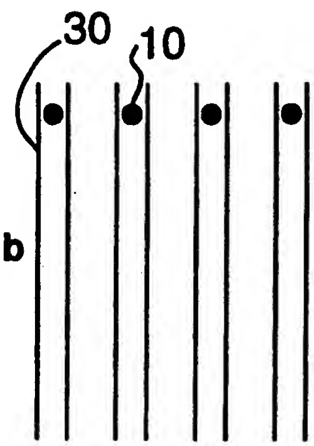

Fig. 1b

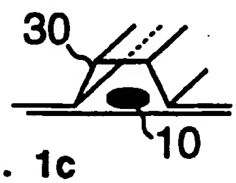

Fig. 1c

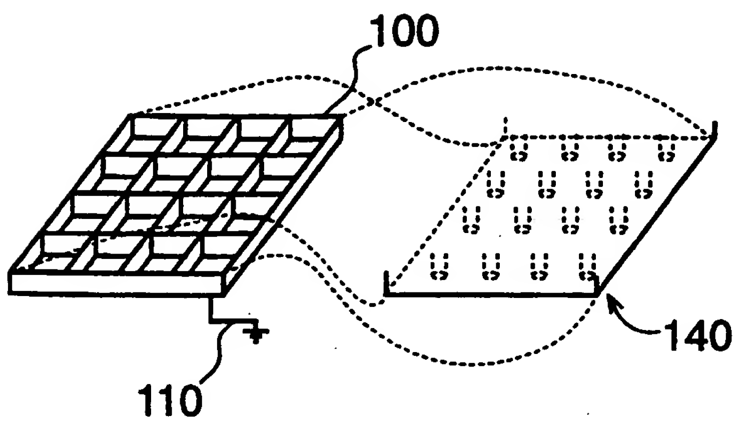

Fig. 2

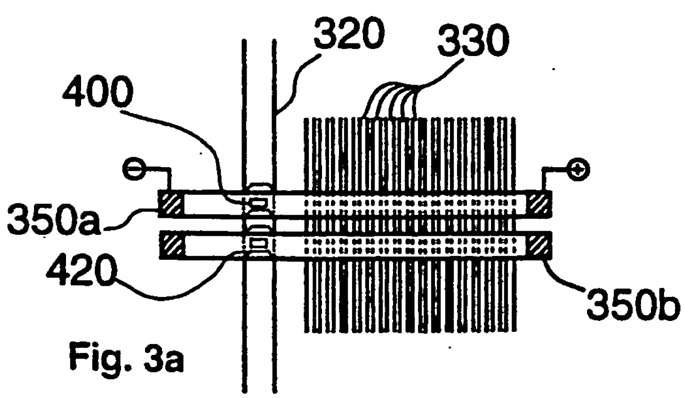

Fig. 3a

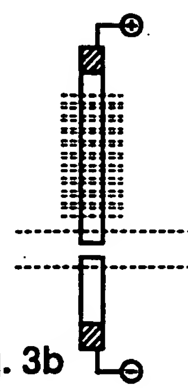

Fig. 3b

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/13313

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :C12Q 1/68; G01N 27/26; C07H 21/00
US CL :435/6; 204/182.8; 536/25.3
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6; 204/182.8; 536/25.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
APS, CAPLUS, MEDLINE, EMBASE, BIOSIS, INPADOC
search terms: Multiplex, sequencing, parallel, polynucleotide, nucleic acid, DNA, array, oligonucleotide, synthesis

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,324,401 A (YEUNG ET AL.) 28 June 1994, columns 1-4, 10, and 16.	7-10
X ---- Y	LAGERKVIST et al. Manifold Sequencing: Efficient Processing of Large Sets of Sequencing Reactions. Proceedings of the National Academy of Sciences USA. March 1994, Vol. 91, pages 2245-2249, especially pages 2245 and 2247.	1-6 ---- 7-10
X	WO 94/05394 A1 (ARRIS PHARMACEUTICAL CORPORATION) 17 March 1994, pages 6-17.	11-16

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
05 DECEMBER 1996	24 DEC 1996

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	AMY ATZEL
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/13313

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-6, drawn to methods and devices for determining sequences.

Group II, claims 7-10, drawn to methods of parallel landing.

Group III, claims 11-16, drawn to methods of synthesis of patterned arrays.

The inventions listed as Groups I-III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Groups I and III do not pertain to capillary electrophoresis as does Group II; and Groups I and II do not require lithographic techniques as does Group III.